Eyewitness Identification Evidence and Improving Legal Safeguards

Alena Skalon

Department of Psychological Sciences
Faculty of Health, Arts & Design
Swinburne University of Technology

This thesis is submitted in total fulfilment of the requirements for the degree of Doctor of Philosophy.

August 2019
Supervisors: Dr. Jennifer Beaudry and Dr. Jordy Kaufman
Abstract

Cases of wrongful convictions, as well as psychological literature, demonstrate that jurors tend to believe eyewitness identifications even when identifications were obtained using suggestive identification procedures and/or were mistaken. The criminal justice system attempts to educate the jury about factors that can influence eyewitness accuracy by providing judicial instructions or expert testimony. This thesis had four research aims. The first aim was to establish the effectiveness of traditional legal safeguards—judicial instructions and expert testimony—in sensitising mock-jurors to the quality of identification procedures and to eyewitness identification accuracy. Taken together, three of my studies (Paper 1 \(N = 232\); Paper 2, Studies 1 \(N = 561\) and 2 \(N = 478\)) suggested that neither of the traditional legal safeguards effectively improved mock-jurors’ decisions. The second aim was to evaluate whether expert testimony’s effectiveness can be improved by accompanying it with visual aids. Contrary to my expectations, expert testimony with visual aids did not outperform traditional expert testimony (Paper 2, Study 1). The third aim was to test whether the type of eyewitness evidence (a video-recorded identification procedure and the eyewitness’s testimony or the eyewitness’s testimony alone) affects mock-jurors’ perceptions of eyewitnesses and the effectiveness of expert testimony. Paper 2, Study 2 demonstrated that mock-jurors who watched eyewitness testimony without the video-recorded identification procedure were more likely to believe that the witness correctly identified the culprit than mock-jurors who watched both. The final aim of the thesis was to examine the effectiveness of an innovative educational aid—the Eyewitness Instructions and Checklist—in sensitising mock-jurors to the quality of identification procedures and to identification accuracy (Paper 3, \(N = 536\) students; \(N = 158\) community members). Paradoxically, the new educational aid led to overbelief: mock-jurors who completed it rendered
more guilty verdicts compared to those in the control condition. Overall, the findings indicate that neither the traditional legal safeguards nor the Eyewitness Instructions and Checklist are effective in improving the accuracy of mock-jurors’ assessments of eyewitness evidence. However, the most consistent finding of this thesis was not directly related to any of the main research aims. In all four studies, a clear pattern emerged: mock-jurors were sensitive to identification accuracy when identification procedures were non-suggestive. Concerningly, this sensitivity was lost when identification procedures were suggestive. This highlights the importance of using best practice procedures when obtaining eyewitness evidence, particularly because the legal safeguards did not sensitise mock-jurors to the damaging effects of suggestive identification procedures on eyewitnesses’ decisions.
Acknowledgements

I would like to express my immense gratitude to my supervisor, Dr. Jennifer Beaudry, for not only believing in me throughout my whole candidature, but also for being understanding and supportive of my ambitions and career goals outside of the university. Without you, Jennifer, I would not be where I am today—not just in terms of my thesis. Thank you for your enthusiastic encouragement, for every supervision session, and for every slice of a PhD dinner pizza.

Thank you to Mehera San Roque for sharing her legal expertise and being open to finding common ground between psychological and legal viewpoints.

I greatly appreciated the assistance provided by the Forensic Psychology Lab students. Thank you for your feedback and your patience in reading my work.

I am profoundly grateful to my husband, Brad Willson, for always being there for me, for cheering me up during many of my moments of self-doubt, for helping me to get through the last few months of my candidature without having a complete mental breakdown. Words cannot express how thankful I am for all your love and support.

Thank you to my friends for being by my side and listening to my complaints, for forgiving my absences, and for providing much-needed distraction from my thesis.

Thank you to my family who always encouraged me to explore the world, be ambitious, and aim high.

My special thanks are extended to Leo Sternbach for his significant contribution to the field of psychopharmacology. I am sure that without his invention the lives of many doctoral students would have been much more difficult.
Declaration

This examinable outcome:

1. Contains no material which has been accepted for the award to the candidate of any other degree or diploma, except where due reference is made in the text of the examinable outcome;

2. To the best of the candidate’s knowledge contains no material previously published or written by another person except where due reference is made in the text of the examinable outcome; and

3. Where the work is based on joint research or publications, discloses the relative contributions of the respective workers or authors.

I warrant that I (the candidate) have obtained, where necessary, permission from the copyright owners to use any third party copyright material reproduced in the thesis (such as artwork, images, unpublished documents), or to use any of my own published work (such as journal articles) in which the copyright is held by another party (such as publisher, co-author).

Signed by Alena Skalon, 5th August 2019
Authorship Indication Forms

Authorship Indication Form
For PhD by Publication candidates

NOTE
This Authorship Indication form is a statement detailing the percentage of the contribution of each author in each published ‘paper’. This form must be signed by each co-author and the Principal Coordinating Supervisor. This form must be added to the publication of your final thesis as an appendix. Please fill out a separate form for each published paper to be included in your thesis.

DECLARATION
We hereby declare our contribution to the publication of the ‘paper’ entitled:

An Interdisciplinary and Cross-National Analysis of Legal Safeguards for Eyewitness Evidence

First Author
Name: Alena Skalon
Signature: ____________________________
Percentage of contribution: 50 %
Date: 02/03/2019
Brief description of contribution to the ‘paper’ and your central responsibilities/role on project:
I completed the literature review and came up with key ideas of this chapter. I wrote the first draft and revised it according to my supervisor’s feedback. I invited Mehera San Roque to be a co-author and further revised and re-worked the chapter.

Second Author
Name: Mehera San Roque
Signature: ____________________________
Percentage of contribution: 10 %
Date: 13/1/19
Brief description of your contribution to the ‘paper’:
My primary role was to review and refine the text dealing with the US and Australian legal frameworks, included in the first substantial draft of the paper, when I was invited to co-author.
This included expanding the Australian and Canadian legal material, but these additional were in line with the paper as conceptualised by Alena.

Third Author
Name: Dr. Jennifer Beaudry
Signature: ____________________________
Percentage of contribution: 10 %
Date: 14/12/2019
Brief description of your contribution to the ‘paper’:
I provided guidance on the direction of the chapter and provided ongoing feedback on the chapter.
Fourth Author

Name: ___________________________ Signature: ___________________________

Percentage of contribution: ____ % Date: ___/___/____

Brief description of your contribution to the "paper":

Principal Coordinating Supervisor:

Name: _______________ Signature: ___________________________

Date: 14/12/2018

In the case of more than four authors please attach another sheet with the names, signatures and contribution of the authors.
Authorship Indication Form
For PhD by Publication candidates

NOTE
This Authorship Indication form is a statement detailing the percentage of the contribution of each author in each published 'paper'. This form must be signed by each co-author and the Principal Coordinating Supervisor. This form must be added to the publication of your final thesis as an appendix. Please fill out a separate form for each published paper to be included in your thesis.

DECLARATION
We hereby declare our contribution to the publication of the 'paper' entitled:

The Effectiveness of Judicial Instructions on Eyewitness Evidence in Sensitizing Jurors to Suggestive Identification Procedures Captured

First Author
Name: Alena Skalon
Signature:
Percentage of contribution: 20% Date: 02/03/2019

Brief description of contribution to the 'paper' and your central responsibilities/role on project:
I assembled stimuli for this project, obtained ethics approval, and programmed the study on Qualtrics.
I recruited participants, monitored study completion, cleaned and analyzed all data. I drafted the manuscript and supervised all aspects following.

Second Author
Name: Dr. Jennifer Beaudry
Signature:
Percentage of contribution: 20% Date: 02/03/2019

Brief description of your contribution to the 'paper':
I assisted with designing the study, provided input into data analysis, and edited various versions of the manuscript.

Third Author
Name:
Signature:
Percentage of contribution: ___% Date: ___/___/____

Brief description of your contribution to the 'paper':
Fourth Author

Name:__________________________Signature:____________________

Percentage of contribution: ____% Date: ___/___/____

Brief description of your contribution to the 'paper':

Principal Coordinating Supervisor:

Name: Jennifer Brindley Signature: ______________________

Date: 07/03/2091

In the case of more than four authors please attach another sheet with the names, signatures and contribution of the authors.

Authorship Indication Form
Authorship Indication Form

For PhD by Publication candidates

NOTE

This Authorship Indication form is a statement detailing the percentage of the contribution of each author in each published ‘paper’. This form must be signed by each co-author and the Principal Coordinating Supervisor. This form must be added to the publication of your final thesis as an appendix. Please fill out a separate form for each published paper to be included in your thesis.

DECLARATION

We hereby declare our contribution to the publication of the ‘paper’ entitled:

The Effects of Expert Testimony on Mock-Surors’ Perceptions of Video-Recorded Eyewitness Identification Evidence

First Author

Name: Alena Skalan
Signature: [Signature]

Percentage of contribution: 60% Date: 03/05/2019

Brief description of contribution to the ‘paper’ and your central responsibilities/role on project:

I created expert testimony stimuli for this project, programmed the study on Qualtrics, and obtained ethics approval. Once ethics clearance was obtained, I recruited participants, analysed all data and wrote the manuscript in consultation with my supervisor.

Second Author

Name: Dr. Jennifer Beattie
Signature: [Signature]

Percentage of contribution: 15% Date: 07/03/2019

Brief description of your contribution to the ‘paper’:

I helped shape the project, provided guidance for data analysis, and provided critical and thoughtful comments on various versions of the manuscript.

Third Author

Name: Travis Edmonds
Signature: [Signature]

Percentage of contribution: 5% Date: 03/05/2019

Brief description of your contribution to the ‘paper’:

I contributed to the programming of Study 1. In addition, I assisted in submitting ethics application for the research.
Fourth Author

Name: __________________________________________ Signature: __________________________

Percentage of contribution: ___%  Date: ____/____/____

Brief description of your contribution to the 'paper':

Principal Coordinating Supervisor:

Name:  Jennifer Brandt  Signature: __________________________

Date: 07/03/2019

In the case of more than four authors please attach another sheet with the names, signatures and contribution of the authors.

Authorship indication Form
Authorship Indication Form
For PhD by Publication candidates

NOTE
This Authorship Indication form is a statement detailing the percentage of the contribution of each author in each published 'paper'. This form must be signed by each co-author and the Principal Coordinating Supervisor. This form must be added to the publication of your final thesis as an appendix. Please fill out a separate form for each published paper to be included in your thesis.

DECLARATION
We hereby declare our contribution to the publication of the 'paper' entitled:

The Eyewitness Instructions and Checklist: An Alternative Method for Educating Surors about Eyewitness Evidence

First Author
Name: Aleda Skalon
Signature: [Signature]
Percentage of contribution: 80% Date: 14/05/2019

Brief description of contribution to the 'paper' and your central responsibilities/role on project:
I came up with the idea of creating a new method created the stimuli, obtained ethics clearance and programmed the study on Qualtrics. I recruited participants and monitored study completion. I cleaned and analyzed the data, drafted the manuscript and made revisions in line with my supervisors feedback.

Second Author
Name: Dr. Jennifer Beadery
Signature: [Signature]
Percentage of contribution: 20% Date: 14/05/2019

Brief description of your contribution to the 'paper':
I helped shape the development of the new method. I provided guidance on the analysis and provided feedback on the manuscript.

Third Author
Name: [Name]
Signature: [Signature]
Percentage of contribution: ___% Date: ___/___/___

Brief description of your contribution to the 'paper':
Fourth Author

Name: ___________________________ Signature: __________________

Percentage of contribution: ___% Date: __/__/____

Brief description of your contribution to the 'paper':

Principal Coordinating Supervisor:

Name: [Signature]

Date: 14/03/2019

In the case of more than four authors please attach another sheet with the names, signatures and contribution of the authors.

Authorship Indication Form
Table of Contents

Abstract .......................................................................................................................... ii
Acknowledgements ........................................................................................................ iv
Declaration ..................................................................................................................... v
Authorship Indication Forms ........................................................................................ vi
List of Tables ................................................................................................................. xvii
List of Figures ............................................................................................................... xix
List of Papers ............................................................................................................... xx

Chapter 1. Overview ..................................................................................................... 1
  Background .................................................................................................................. 1
  The Development of the Focus of the Research ......................................................... 6
  A Thematic Overview of Papers ............................................................................... 8
  Research Aims ............................................................................................................ 9
  Outline of Papers ...................................................................................................... 10
  Implications ............................................................................................................... 17

Chapter 2. Literature Review ...................................................................................... 18
  Wrongful Convictions, Eyewitness Evidence, and Jurors’ Decision-Making .......... 20
  Admissibility Safeguards ......................................................................................... 31
  Corrective Safeguards ............................................................................................... 39
  Comparative Analysis: Judicial Instructions vs. Expert Testimony ....................... 62
  Conclusions and Future Directions ........................................................................ 66

Chapter 3. Method ....................................................................................................... 71

Chapter 4. Paper 1 ....................................................................................................... 81
  Abstract .................................................................................................................... 82
  Introduction .............................................................................................................. 83
  Method ...................................................................................................................... 89
  Results ...................................................................................................................... 93
  Discussion ............................................................................................................... 103

Chapter 5. Paper 2 ....................................................................................................... 109
  Abstract ................................................................................................................... 110
  Introduction ............................................................................................................. 112
Study 1 .............................................................................................................................. 117
Study 2 .............................................................................................................................. 135
General Discussion ........................................................................................................ 147
Supplemental Material ..................................................................................................... 154

Chapter 6. Paper 3 ........................................................................................................... 157
Abstract ........................................................................................................................... 158
Introduction ...................................................................................................................... 160
Method .............................................................................................................................. 166
Results .............................................................................................................................. 171
Discussion ......................................................................................................................... 180
Supplementary Materials .................................................................................................. 185

Chapter 7. Conclusion ..................................................................................................... 198
Synthesis of Findings ....................................................................................................... 201
Lack of a Theoretical Model ............................................................................................ 207
Incorporating Psychological Findings into Legal Practice ............................................. 209
Limitations ....................................................................................................................... 212
Future Research ............................................................................................................... 213
Conclusion ......................................................................................................................... 216

References ........................................................................................................................ 218
List of Appendices

Appendix 2. Victorian Judicial Instructions .................................................................................................273
Appendix 3. Full Trial Transcript used in Paper 2 .......................................................................................276
Appendix 4. PowerPoint Slides that Accompanied Expert Testimony in Paper 2 ..................286
Appendix 5. The Eyewitness Instructions and Checklist Transcript .........................................................292
Appendix 6. The Eyewitness Instructions and Checklist (Checklist) .........................................................298
Appendix 7. Case Summary and Attorneys’ Arguments used in Paper 3 ..................................................300
Appendix 8. Dependent Variables ............................................................................................................301
Appendix 9. Ethics Declaration ..................................................................................................................303
Appendix 10. Ethics Approval Statements .................................................................................................304
Appendix 11. Ethics Final Reports ............................................................................................................308
Appendix 12. List of Presentations and Publications .................................................................................310
**List of Tables**

Table 1. Witnessing and Identification Conditions.................................................................25
Table 2. Summary of Studies that Tested the Effectiveness of Judicial Instructions...............49
Table 3. Summary of Studies that Tested the Effectiveness of Expert Testimony......................60
Table 4. Summary of Studies that Compared Judicial Instructions and Expert Testimony .......64
Table 5. Design Summary...........................................................................................................72
Table 6. Questions Assessing the Comprehension of the Judicial Instructions, and Percentage of Mock-jurors who Gave the Correct Answer (Paper 1) ..................................................................................102
Table 7. Mock-jurors’ Mean Ratings for the Continuous Dependent Variables as a Function of Expert Testimony, Identification Accuracy, and Identification Procedure (Paper 2, Study 1) ...127
Table 8. Percentage of Correct Answers in Control and Expert Testimony Conditions (Paper 2, Study 1) ........................................................................................................................................132
Table 9. Mock-jurors’ Mean Ratings for the Likelihood that the Eyewitness Correctly Identified the Culprit as a Function of Expert Testimony, Identification Accuracy, and Type of Evidence (Paper 2, Study 2) ........................................................................................................................................140
Table 10. Mock-Jurors’ Mean Ratings for the Perceptions of the Administrator According to the Two-Way Interactions between ID Procedure and Expert Testimony, Type of Evidence, and ID Accuracy (Paper 2, Study 2) ....................................................................................................................................142
Table 11. Mock-Jurors’ Perceptions of the Eyewitness According to the Two-Way Interactions of ID Accuracy with Expert Testimony and ID Procedure (Paper 2, Study 2) .......................144
Table 12. Percentage of Correct Answers in Control and Expert Testimony Conditions (Paper 2, Study 2) ........................................................................................................................................145
Table 13. Percentage of Guilty Verdicts as a Function of Expert Testimony, Identification Accuracy, Type of Evidence, and Identification Procedure (Paper 2, Study 2) .......................155
Table 14. Mock-jurors’ Mean Ratings for Confidence-in-Verdict as a Function of Expert Testimony, Identification Accuracy, and Identification Procedure (Paper 2, Study 2) .............156
Table 15. Univariate Analyses of Variance Results for ID Procedure × ID Accuracy × Educational Aid Analyses on Continuous Dependent Variables (Paper 3).................................175
Table 16. Mock-jurors’ Mean Ratings (with Standard Deviations) for the Continuous Dependent Variables as a Function of Educational Aid, Identification Accuracy, and Identification Procedure (Paper 3).........................................................................................................................176
Table 17. Percentage of Correct Answers to Knowledge Questions in Control and the EIC Conditions (Student Sample; Paper 3)........................................................................................................................................179
Table 18. Binary Logistic Regression Analyses Predicting Mock-jurors’ Verdicts from Sample Type and Identification Accuracy, Identification Procedure, Educational Aid (Paper 3) ..........186
Table 19. Hierarchical Logistic Regression Model Building Summary for Students’ Verdicts (Paper 3)................................................................................................................................................187
Table 20. Hierarchical Logistic Regression Model Building Summary for Community Members’ Verdicts (Paper 3)........................................................................................................................................188
Table 21. Univariate Analyses of Variance Results for ID Procedure × ID Accuracy × Educational Aid Analyses on Continuous Dependent Variables for Community Members (Paper 3) ..................................................................................................................................................190
Table 22. Community Members’ Mean Ratings (with Standard Deviations) for the Continuous Dependent Variables as a Function of Educational Aid, Identification Accuracy, and Identification Procedure (Paper 3)..................................................................................................................................................191
Table 23. Percentage of Correct Answers to Knowledge Questions by Control and the EIC Conditions (Community Members Sample; Paper 3)......................................................................................................................194
Table 24. Bayesian Models with the Highest Bayes Factor for Students and Community Members (Paper 3)..................................................................................................................................................197
List of Figures

Figure 1. Sample Lineups ........................................................................................................73
Figure 2. General Outline of the Procedure .............................................................................74
Figure 3. Still Images of an Identification Procedure and the Eyewitness’s Testimony .......75
Figure 4. A Slide Used in Expert Testimony with Visual Aids Condition, Accompanying Verbal Information about the Post-Identification Feedback........................................................................77
Figure 5. Part of the Checklist from the EIC ........................................................................78
Figure 6. Two-way interaction between identification accuracy and identification procedure on mock-jurors’ verdicts (Paper 1) .........................................................................................................................95
Figure 7. Two-way interaction between identification accuracy and identification procedure on mock-jurors’ confidence-in-verdict (Paper 1) .........................................................................................................................97
Figure 8. Two-way interaction between identification procedure and judicial instructions on mock-jurors’ estimate of the likelihood that the witness made an accurate identification (Paper 1) .........................................................................................................................98
Figure 9. Two sample slides that accompanied the expert testimony in the expert testimony with visual aids condition (Paper 2) ..........................................................................................................................121
Figure 10. The two-way interaction between identification accuracy and identification procedure on mock-jurors’ verdicts (Paper 2) .........................................................................................................................125
List of Papers


Chapter 1. Overview

Background

Eyewitness evidence is involved in approximately 70% of wrongful convictions (Innocence Project, 2019). It is important to note that a mistake made by an eyewitness does not directly lead to a conviction—it is up to the judge or the jury to evaluate the evidence and decide how much weight to give to it. The problem, however, is that jurors are generally unaware of factors that influence the reliability and accuracy of eyewitness identifications. The criminal justice system attempts to resolve this issue by providing information about the psychology of eyewitness identification either in the form of judicial instructions or in the form of expert testimony. Empirical research demonstrated that neither of these legal safeguards (i.e., judicial instructions and expert testimony) effectively improve mock-jurors’ ability to differentiate between weak and strong eyewitness evidence (e.g., Devenport & Cutler, 2004; Jones, Bergold, Dillon, & Penrod, 2017; Martire & Kemp, 2009). Little is known, however, about the effectiveness of legal safeguards in improving jurors’ sensitivity to identification conditions and to eyewitness accuracy when jurors are presented with genuine eyewitnesses (people who actually saw a crime) rather than simulated testimony. Some scholars contend that alternative methods of educating the jury may be more effective than current traditional legal safeguards (e.g., Jones et al., 2017; Mnookin, 2015). Moreover, some states, researchers, and national bodies (e.g., National Academy of Sciences, USA) recommended video-recording the initial identification procedures. Despite the presumed value of doing so, the question remains as to whether these video recordings should be presented to jurors in all situations and whether this type of evidence would influence the effectiveness of legal safeguards.
**Jurors’ perceptions of eyewitness evidence.** Eyewitnesses often play a crucial role in trial processes because they can shed light on one of the main issues in criminal cases—the identity of the perpetrator. Despite the utility of eyewitness testimony, real cases and psychological research demonstrated that eyewitnesses sometimes misidentify innocent people as perpetrators (e.g., The Innocence Project, 2019). Furthermore, people often believe eyewitnesses regardless of the actual quality of the evidence (e.g., the witness’s exposure to the perpetrator or how the identification was conducted; for an overview, see Boyce, Beaudry, & Lindsay, 2007). One of the possible explanations for this tendency to believe eyewitnesses is jurors’ lack of knowledge about the factors that influence the reliability of eyewitness evidence (e.g., Benton, Ross, Bradshaw, Thomas, & Bradshaw, 2006).

**The effectiveness of legal safeguards.** Legal safeguards can sometimes be used to educate jurors about the psychology of eyewitness memory. The traditional legal safeguards include judicial instructions and expert testimony. Judicial instructions are read by the judge; they provide information about eyewitness factors and prompt jurors to evaluate the quality of eyewitness evidence. Expert testimony, on the other hand, is delivered by a psychologist who is examined by one lawyer (usually the defence) and is subject to cross-examination by the other lawyer. Expert testimony can include general information about the psychology of memory and present findings of psychological research (for a detailed discussion see Chapter 2).

Both safeguards can produce one of four outcomes on the jurors’ decisions: a null effect, confusion, skepticism towards eyewitness evidence, or sensitivity to the quality of eyewitness evidence (e.g., Cutler, Penrod, & Dexter, 1989; Martire & Kemp, 2011). A null effect would occur if a legal safeguard did not significantly influence jurors’ assessments of eyewitness evidence. Confusion refers to a situation when jurors evaluate eyewitness evidence contrary to
the content of a legal safeguard (e.g., more guilty verdicts when witnessing conditions were poor). Skepticism towards eyewitness evidence would be evident if jurors discount an eyewitness identification regardless of the strength of the evidence (e.g., not believing a correct identification). Finally, the only desirable outcome of a legal safeguard is sensitivity to the quality of eyewitness evidence. An effective legal safeguard would sensitise jurors to either (a) witnessing and identification conditions or (b) identification accuracy. Jurors who are sensitive to witnessing and identification conditions would believe eyewitnesses who had good witnessing conditions (e.g., good lighting) and whose identifications were obtained using non-suggestive identification procedures (e.g., a fair lineup) more often than those who witnessed a crime under poor conditions (e.g., brief exposure to the culprit’s face) and identified the suspect under suggestive circumstances (e.g., showup). Similarly, jurors who are sensitive to identification accuracy would believe correct identifications and discount mistaken ones (see Chapter 2 for a more detailed discussion of these effects).

The majority of studies that tested the effectiveness of legal safeguards presented participants with simulated eyewitness evidence—that is, eyewitnesses were played by actors. These studies equated good witnessing (e.g., good lighting) and identification conditions (e.g., non-suggestive lineup instructions) with reliable eyewitness evidence. Generally, an eyewitness who had good witnessing conditions and made an identification from a non-suggestive procedure is more likely to be correct than an eyewitness who saw the perpetrator in poor witnessing conditions and identified him/her from a suggestive procedure. Nonetheless, good witnessing and identification conditions should not be equated with eyewitness accuracy. In a real trial, jurors have to decide if the eyewitness correctly identified the culprit or misidentified an innocent person, taking into account multiple factors, including the quality of witnessing and
identification conditions, verbal and non-verbal behaviour of the eyewitness, and the eyewitness’s confidence. However, in studies that simulated eyewitness evidence the latter two factors—verbal and non-verbal behaviour and genuine confidence statement—are not present, depriving participants of this crucial information. A genuine eyewitness paradigm, on the other hand, presents participants with videos of genuine eyewitnesses rather than eyewitness-actors who read a script. In this paradigm, eyewitness-participants first witness a crime and are then asked to identify the culprit and testify about their experience (Wells, Lindsay, & Ferguson, 1979). This paradigm is critical to shed light on the effectiveness of legal safeguards because it helps to establish the ultimate accuracy of jurors’ judgments. Because the stimuli are produced in a laboratory, the researchers know the ground truth—whether the eyewitness is correct or mistaken.

Empirical tests of judicial instructions demonstrate that they most often do not affect mock-jurors’ judgments (e.g., Cutler, Dexter, & Penrod, 1990; Jones et al., 2017; Martire & Kemp, 2009) or lead to skepticism (e.g., Dillon, Jones, Bergold, Hui, & Penrod, 2017; Papailiou, Yokum, & Robertson, 2015). Expert testimony either leads to skepticism (e.g., Cutler et al., 1990; Jones et al., 2017; Leippe, Eisenstadt, Rauch, & Seib, 2004), or to limited sensitivity (e.g., Cutler, Dexter, & Penrod, 1989; Devenport et al., 2002; Geiselman & Mendez, 2005). For example, Cutler, Dexter, and Penrod (1989) found that expert testimony sensitised mock-jurors’ to witnessing and identification conditions, however, this sensitivity occurred only for the strength of the prosecution’s case and of the defence’s case; there was no effect on the ultimate decision—the verdict. Importantly, only two studies, to my knowledge, used a genuine eyewitness paradigm—both demonstrated the ineffectiveness of judicial instructions and expert
testimony in sensitising mock-jurors to identification accuracy (Wells, Lindsay, & Tousignant, 1980; Martire & Kemp, 2009).

**Educating the jury.** Given the overwhelming evidence that the traditional safeguards are ineffective, traditional approaches to jury education should be reconsidered. For example, the addition of visual aids to verbal expert testimony improved mock-jurors’ understanding and application of the information provided by an expert (Jones & Kovera, 2015; Park & Feigenson, 2013). Furthermore, novel methods of educating jurors may be more effective than traditional legal safeguards. For example, Wise, Fishman, and Safer (2009) created the interview-identification-eyewitness (I-I-Eye) educational tool. The I-I-Eye presents PowerPoint slides that educate jurors about eyewitness factors. The I-I-Eye successfully sensitised jurors to the quality of eyewitness evidence: jurors were more likely to vote guilty when the case was strong, compared to weak (cf. jury duty aid and Neil v. Biggers aid [Neil v. Biggers, 1972]; Pawlenko, Safer, Wise, & Holfeld, 2013; Safer et al., 2016). The I-I-Eye has not yet been tested with genuine eyewitness evidence, thus, it is unclear if it would improve jurors’ sensitivity to identification accuracy. Alternative methods of educating jurors might be more effective than traditional legal safeguards because they can simplify the presentation of information and provide additional tools that may assist jurors in applying this information to the case at hand. More research is needed, however, to decidedly establish the effectiveness of alternative educational aids, particularly with genuine eyewitnesses and with different types of eyewitness evidence (e.g., video-recorded identification procedures).

**Video-recorded identification procedures.** Another method that may assist jurors in their evaluations of eyewitness evidence is the presentation of a video-recorded identification procedure. This video may provide useful information about the confidence of an eyewitness and
preserve the behavioural cues that are otherwise lost during the eyewitness’s preparation for a trial. Indeed, when mock-jurors were exposed to a video-recorded identification procedure along with the eyewitness’s testimony, they were better able to differentiate between correct and mistaken identifications (Reardon & Fisher, 2011) and between non-suggestive and suggestive identification procedures (Mojadidi & Kovera, 2018) compared to participants who viewed the eyewitness’s testimony alone. However, this effect was not observed when identifications were obtained under suggestive circumstances (Beaudry et al., 2015), which mimicked the results found with testimony alone (Smalarz & Wells, 2014). These results indicate that jurors may be able to discriminate between correct and mistaken identifications, but only when they are obtained using non-suggestive identification procedures. No published research to date has examined how the presentation of video-recorded identification procedures from genuine eyewitnesses would interact with the effectiveness of legal safeguards. A video-recorded identification procedure is a very different type of evidence compared to an eyewitness testimony. As more jurisdictions video-record identification procedures, this new type of evidence will likely be presented to jurors. It is essential to understand whether legal safeguards are effective with this new type of evidence.

The Development of the Focus of the Research

The initial focus of my project was on exploring the effectiveness of various types of expert testimony and combining expert testimony with a video-recorded identification procedure. Typically, expert testimony on the psychology of eyewitness identification includes a verbal overview of general findings of eyewitness literature. The initially proposed program of research included three studies. The first study would have investigated whether visual aids improve mock-jurors’ understanding of eyewitness evidence and sensitise them to suggestive procedures.
Thus, it would have compared verbal general expert testimony with and without visual aids (such as graphs and bar charts). The second study would have focused on the issue of specificity and compared general and specific expert testimony, both accompanied with visual aids. General expert testimony would have presented the overview of findings of eyewitness literature, whereas specific expert testimony would have discussed the presence of each factor in the particular case. Finally, the third study would have investigated whether the timing of expert testimony (i.e., before or after eyewitness evidence) influences its effectiveness. In trials, the expert usually testifies after the eyewitness, which can potentially limit the effectiveness of the expert’s advice; by the time jurors hear it, they may not be able to retrospectively apply the information provided by the expert to the case at hand (Pennington & Hastie, 1986).

However, at my confirmation of candidature review, panel members indicated that my research proposal did not contain any information about another commonly used legal safeguard—judicial instructions. Given that Australian courts rarely rely on expert testimony, I decided that it was important to examine the most commonly used safeguard in our state of Victoria. No empirical work has examined the effectiveness of Victorian judicial instructions in eyewitness cases; thus, my work in this area both expands our current knowledge of the effectiveness of judicial instructions and provides specific information about the legal safeguard regularly relied upon in Victorian courts. Furthermore, given the overwhelming evidence of the ineffectiveness of traditional legal safeguards, I created a new method of educating the jury—the Eyewitness Instructions and Checklist (EIC). Thus, the focus of my project shifted from exploring only the effectiveness of expert testimony to exploring the effectiveness of legal safeguards more broadly.
I conducted four studies. All of my studies employed a genuine eyewitness paradigm and presented mock-jurors with a video-recorded identification procedure along with the eyewitness’s testimony (except Paper 2, Study 2 in which I manipulated the type of evidence). The first study tested the effectiveness of the Victorian judicial instructions, the second and third tested the effectiveness of expert testimony, and the fourth study explored the effectiveness of the EIC.

In addition to conducting my empirical work, I also sought to examine a recent legal decision, Bayley v The Queen (2016) from a psychological perspective. That manuscript is not included as part of my thesis, however, it reviews the issues with jurors’ perceptions of new types of eyewitness evidence. The victim made the initial identification from a photograph seen on Facebook and then later in a formal lineup procedure. The jury convicted the defendant despite multiple weaknesses present in the case. The Court of Appeal, on the other hand, argued that the identification evidence should not have been permitted at the initial trial. I reviewed the Court’s reasoning in light of empirical evidence and discussed the implications of emerging types of evidence for legal practice and for psychological research (see Appendix 1).

A Thematic Overview of Papers

This thesis includes seven chapters. The current—first—chapter is introductory and provides an overview of my thesis. The second chapter serves as my literature review and provides an overview of issues related to legal safeguards. The third chapter gives a detailed description of the method and materials used in this thesis. Following it, I present three empirical chapters. Chapter 4 (Paper 1) presents the results of the study testing the effectiveness of judicial instructions. Chapter 5 (Paper 2) discusses two studies investigating the effectiveness of traditional expert testimony, expert testimony with visual aids, and exploring whether the type of
evidence (a video-recorded identification procedure and the eyewitness’s testimony compared to the eyewitness’s testimony alone) influences the effectiveness of expert testimony. Chapter 6 (Paper 3) presents the results of the final study investigating the effectiveness of the Eyewitness Instructions and Checklist. Chapter 7 discusses these four studies and considers their implications for future research.

Research Aims

Research aim one. Traditionally courts have relied on two legal safeguards to educate jurors about the psychology of eyewitness identifications—judicial instructions and expert testimony. Despite widespread reliance on these safeguards (see Chapter 2 for a review), empirical research has produced mixed results regarding their effectiveness. Moreover, a very limited number of studies utilised a genuine eyewitness paradigm. Thus, the first research aim of this thesis was to establish the effectiveness of judicial instructions and expert testimony in improving the accuracy of mock-jurors’ assessments of genuine eyewitnesses (findings described in Chapters 4 and 5).

Research aim two. In addition to testing traditional legal safeguards, I evaluated whether accompanying traditional expert testimony with visual aids would improve its effectiveness. I created PowerPoint slides that presented a text summary of eyewitness factors and graphs that were designed to draw mock-jurors’ attention to key points of expert testimony (see Chapter 5: Study 1).

Research aim three. Eyewitness evidence is typically presented in the form of in-court testimony in which the eyewitness outlines her or his memory of the event in question. All studies of traditional legal safeguards have, to my knowledge, utilised eyewitness testimony only (simulated or genuine), whereas in my thesis mock-jurors watched a video-recorded
identification procedure as well as the eyewitness’s testimony. Despite some benefits of this approach, this methodological consideration also limited my ability to make direct comparisons between my work and previous empirical research on the topic. In order to clearly separate the effects of legal safeguards from the effects of the type of evidence, I conducted a study comparing the effectiveness of expert testimony using both types of evidence (a video-recorded identification procedure and the eyewitness’s testimony vs. the eyewitness’s testimony only; see Chapter 5: Study 2).

**Research aim four.** Given the mixed results produced by studies that tested the effectiveness of traditional legal safeguards, I created an innovative method of educating jurors—the Eyewitness Instructions and Checklist. Thus, the final aim of this thesis was to evaluate whether the EIC is effective in sensitising mock-jurors to the quality of identification procedures and to identification accuracy (see Chapter 6).

**Outline of Papers**

**Literature review.** The literature review provided an interdisciplinary and cross-national perspective on the effectiveness of the two most commonly-used legal safeguards—judicial instructions and expert testimony. In this paper, my co-authors and I provide the background information necessary to understand the importance of using effective legal safeguards before delving into the discussion of the empirical effectiveness of the two legal safeguards. We then compare the evidentiary rules and procedures used in Australia, Canada, and the United States of America and discuss whether these rules and procedures are supported by empirical evidence. Finally, I offer a comparative analysis of judicial instructions and expert testimony by integrating the legal and psychological perspectives. This paper is an important connection between empirical studies and real-world demands and issues faced by lawyers and judges who work with
eyewitnesses. Our invited chapter proposal was accepted and the chapter will become a part of the “Advances in Psychology and Law” series, published by Springer. The editors have reviewed an earlier version of the chapter on the 19th of September, 2018 and requested no substantive changes.

**Scholarly contribution.** The main contribution of this paper is that it bridges psychologists and legal scholars interested in these legal safeguards and offers both perspectives in one chapter. To my knowledge, it is the first cross-national comparison of admissibility safeguards and corrective trial safeguards in Australia, Canada, and the USA.

**My contribution.** I completed the literature review and developed the key ideas for this chapter. I wrote the first draft and revised it in line with my supervisor’s feedback. It soon became clear that a legal perspective would strengthen the paper, thus, I invited Mehera San Roque (senior lecturer in the university of New South Wales, faculty of law) to co-author this chapter and reworked the chapter incorporating her feedback and legal expertise. My contribution constitutes approximately 80% of the work for this paper.

**Paper 1.** The first empirical paper tested the effectiveness of judicial instructions. The study evaluated whether the judicial instructions used in Victoria, Australia improve mock-jurors’ ability to differentiate between suggestive vs. non-suggestive identification conditions and correct vs. mistaken identifications. This was the first study, to my knowledge, that tested whether judicial instructions improve mock-jurors’ sensitivity to video-recorded genuine eyewitness evidence. Mock-jurors ($N = 224$) read a trial transcript and watched a video of an eyewitness’s identification and the eyewitness’s testimony. I varied the suggestiveness of the identification procedure, identification accuracy, and the presence of judicial instructions. The results indicated that judicial instructions did not affect mock-jurors’ verdicts. This was in line
with the results of studies that tested the effectiveness of judicial instructions in the USA (Jones et al., 2017) and Australia (Martire & Kemp, 2009). This study demonstrated that judicial instructions did not affect mock-jurors’ judgments even when they could see how the lineup was administered and watch the eyewitness make their identification.

**Scholarly contribution.** In conversation with Victorian lawyers, they expressed the view that Victorian judicial instructions are superior to those used in other jurisdictions. This was the first study that tested Victorian judicial instructions, providing empirical evidence of their ineffectiveness. This is also a significant contribution to the international literature on judicial instructions because the majority of tests were carried out using American judicial instructions (i.e., Telfaire, Henderson), with the exception of Martire and Kemp (2009) who tested the judicial instructions used in New South Wales, Australia. My study strengthened the conclusion that despite differences in the content of the instructions, judicial instructions are an ineffective legal safeguard against risks associated with eyewitness misidentification.

**My contribution.** I assembled the stimuli for this study which included finding the Victorian judicial instructions, editing and uploading eyewitness videos, and writing the case transcript (see Appendix 2 for the full text of Victorian judicial instructions). Once I prepared all materials, I obtained ethics clearance, designed and programmed the study on Qualtrics, and orchestrated several rounds of test completions. I recruited participants, monitored study completion, and assigned correct course credits in exchange for their participation (in accordance with the Swinburne Psychology Research Experience Program). Prior to analysing the full data set I registered the study on the Open Science Framework using the AsPredicted registration form. I then cleaned and analysed all data and wrote the first draft of the paper which I revised incorporating my supervisor’s feedback, and submitted it to the *Journal of Experimental*
Criminology. I also uploaded materials and both raw and clean data sets to the Open Science Framework. The editorial board recognised the contribution of this paper and the manuscript received a “revise and resubmit” decision on the 22nd of February, 2019. My contribution constitutes approximately 80% of the work for this paper.

**Paper 2.** The second empirical paper includes two studies. Study 1 (N = 561) examined the effectiveness of traditional expert testimony and expert testimony with visual aids. Study 2 (N = 478) tested whether the type of evidence (a video-recorded identification procedure and the eyewitness’s testimony or the eyewitness’s testimony only) influence the effectiveness of expert testimony. In both studies all mock-jurors watched a case summary read by the judge, as well as attorneys’ opening and closing arguments. The main piece of evidence was eyewitness identification which varied in terms of identification accuracy and the suggestiveness of the identification procedures. After watching the eyewitness evidence, mock-jurors either viewed traditional expert testimony, expert testimony with visual aids (Study 1 only), or completed a filler activity (control). In Study 2, I also manipulated the type of evidence. Both studies demonstrated that expert testimony is not effective in sensitising mock-jurors to the quality of identification procedures or to identification accuracy. Moreover, there was some evidence that expert testimony led to skepticism. Expert testimony with visual aids did not outperform traditional expert testimony on any of the measures. Study 2 demonstrated that mock-jurors were more willing to believe that the eyewitness correctly identified the culprit when they watched the eyewitness’s testimony compared to a video-recorded identification and the eyewitness’s testimony. Furthermore, these studies demonstrated that mock-jurors were sensitive to identification accuracy when identification procedures were non-suggestive, but sensitivity was lost when identification procedures were suggestive.
**Scholarly contribution.** These were the first studies, to my knowledge, that tested the effectiveness of expert testimony when combined with a video-recorded identification procedure and the testimony of a genuine eyewitness. Expert testimony is one of the most popular legal safeguards, however, my studies further demonstrated its ineffectiveness. Furthermore, Study 2 contributed to the very limited literature examining mock-jurors’ perceptions of video-recorded identification procedures and highlighted that it is a fundamentally different type of evidence from eyewitness testimony and more research is needed to establish whether video-recorded identification procedures should be made available to jurors. Finally, the damaging effects of suggestive identification procedures uncovered in these studies underscore the importance of using best practice procedures when obtaining eyewitness identifications.

**My contribution.** My contribution was similar to that outlined in Paper 1. I created the expert testimony stimuli based on transcripts of real-life expert testimonies, found actors to play the judge, attorneys, and the expert (see Appendix 3 for the full trial transcript and Appendix 4 for PowerPoint slides). I organised the video-recording, editing, and uploading of expert testimony. In addition, I edited and uploaded the eyewitness videos. I obtained ethics approval, programmed Study 1 on Qualtrics, and recruited participants. Travis Edmonds assisted in the submission of the ethics application for Study 2 and programmed the study on Qualtrics. Travis used a subset of the data for his thesis. I continued to collect additional data and the results reported in the paper are based on the final sample. Once data collection was complete, I cleaned and analysed data for both studies and wrote the first draft of the paper. I also registered the study and uploaded materials and data sets to the Open Science Framework. Following ongoing feedback from my supervisor, I revised the paper to improve readability and added tables to simplify the presentation of statistical information. Finally, I submitted the paper to *Law and*
Human Behavior on the 6th of March 2019. The manuscript is currently under review. My contribution constitutes approximately 80% of the work for this paper.

**Paper 3.** The third—and final—paper investigated whether the alternative method of educating jurors (i.e., the Eyewitness Instructions and Checklist [EIC]) can sensitise mock-jurors to the suggestiveness of identification procedures and to identification accuracy. As demonstrated in Papers 2 and 3, traditional legal safeguards were not effective in improving the accuracy of mock-jurors’ evaluations of eyewitness evidence. For this reason, I created the EIC. The EIC includes several innovative features, such as educating jurors before the trial, presenting information about the psychology of eyewitness identification in a video-recorded format, and providing a checklist that can help jurors integrate the information to the case at hand and, thus, correctly evaluate the quality of the evidence. Mock-jurors ($N = 536$ students; $N = 158$ community members) either watched the EIC or a jury duty video (control condition). They then read the case summary, attorneys’ opening statements, and watched the eyewitness evidence. Similarly to my other studies, I manipulated the suggestiveness of the identification procedures and accuracy of the identification. Following the presentation of the evidence, mock-jurors either completed the second part of the EIC (the checklist) or completed a control activity, rendered a verdict, and completed a questionnaire. The EIC did not improve mock-jurors’ sensitivity to the quality of identification conditions or to identification accuracy. Paradoxically, the EIC led to overbelief, which is in contrast to all previous research on the effectiveness of legal safeguards. Similarly to the findings of my other studies, mock-jurors were sensitive to identification accuracy when identification procedures were non-suggestive, but not when they were suggestive.
Scholarly contribution. The ineffectiveness of traditional legal safeguards warrants the investigation of alternative methods of educating jurors about eyewitness evidence. This study advanced our knowledge of the effectiveness of alternative methods; to date, researchers have developed and tested only one alternative method (i.e., the interview-identification-eyewitness educational aid, Wise et al., 2009). Two key differences between the interview-identification-eyewitness educational aid and the EIC is the timing of presentation and the presence of a checklist. The checklist might be responsible for producing overbelief; thus, future attempts to create alternative methods of educating jurors should test its effects by systematically manipulating its use.

My contribution. For this study, in addition to performing the same tasks outlined in my contribution statements for Papers 1 and 2, I created the EIC. I reviewed the eyewitness literature, various judicial instructions, and transcripts of expert testimonies to identify eyewitness factors that should be discussed in the EIC. I also reviewed the literature describing the use of checklists in medicine (e.g., Hales, Terblanche, Fowler, & Sibbald, 2007) prior to constructing the EIC checklist (see Appendix 5 for the EIC transcript and Appendix 6 for the EIC checklist). Once I created all necessary materials, I obtained ethics approval and recruited student participants as mock-jurors. I also decided to recruit community members to compare the performance of students and community members. I obtained ethics approval and applied for and won a grant from the Association for Psychological Science (USA) Student Grant Competition. Using this grant and the university-allocated research funds I recruited community members. Once all data was collected, I wrote the paper and, after incorporating my supervisor’s feedback, submitted it to the Journal of Applied Research in Memory and Cognition on the 15th of March. Similarly to Papers 1 and 2, I registered the study (prior to the analysis of the full data set) and
uploaded materials and data to the Open Science Framework. My contribution constitutes approximately 80% of the work for this paper.

**Implications**

This thesis will advance our knowledge of the methods that can help jurors comprehend trial-relevant information and improve the accuracy of their judgments in cases that involve eyewitness identification evidence. Findings of the studies included in this thesis might assist in preventing wrongful convictions based on eyewitness misidentifications by empirically establishing the effectiveness of legal safeguards that are currently used in courts. Most of the previous research that tested the effectiveness of traditional legal safeguards used simulated eyewitness evidence. A few studies that used a genuine eyewitness paradigm did not manipulate the suggestiveness of identification procedures (Wells, Lindsay, & Tousignant, 1980; Martire & Kemp, 2009). Prior to my work, there were no studies that tested the effectiveness of legal safeguards in sensitising mock-jurors to eyewitness accuracy and to the quality of identification conditions when they view genuine eyewitness evidence. In addition, technological developments can change the way evidence is presented in court. For example, an initial identification procedure can be video-recorded and presented in front of the jury. Research into jurors’ perceptions of such evidence is slowly accumulating and no studies have yet examined how this type of evidence interacts with legal safeguards. This project aims to fill these knowledge gaps.
Chapter 2. Literature Review

An Interdisciplinary and Cross-national Analysis of Legal Safeguards for Eyewitness Evidence

Alena Skalon, Mehera San Roque, & Jennifer L. Beaudry

Author note: Alena Skalon, PhD Candidate, Department of Psychological Sciences; School of Health Sciences; Faculty of Health, Arts and Design; Swinburne University of Technology, Hawthorn, Victoria, Australia; email: askalon@swin.edu.au.

Mehera San Roque, Senior Lecturer, Faculty of Law; University of New South Wales, New South Wales, Australia; email: m.sanroque@unsw.edu.au.

Dr. Jennifer L. Beaudry, Senior Lecturer, Department of Psychological Sciences; School of Health Sciences; Faculty of Health, Arts and Design; Swinburne University of Technology, Hawthorn, Victoria, Australia; email: jbeaudry@swin.edu.au.

Please note that I have made small alterations to this Paper for thesis presentation. These changes were limited to placement of tables and figures, numbering of tables and figures, removing all reference lists (which are compiled at the end of the thesis), and altering some in-text citations to align with this compiled list.
Abstract

Traditional legal safeguards, along with evidentiary rules and procedures, are intended to ensure a fair trial for criminal defendants and to avoid the conviction of the innocent. In this chapter, we focus on admissibility safeguards and two corrective trial safeguards that provide decision-makers with information about the reliability of eyewitness identification evidence—judicial instructions and expert testimony. We briefly provide background information about wrongful convictions and the science of eyewitness evidence to give context to the examination of the effectiveness of legal safeguards in cases with eyewitness identification evidence. We then present a cross-national comparison of admissibility safeguards and corrective trial safeguards in Australia, Canada, and the USA. We also review the empirical studies that have tested the effectiveness of judicial instructions and expert testimony. Based on this interdisciplinary analysis, we conclude that courts are too permissive in admitting eyewitness evidence and are generally too reliant on the corrective safeguards, despite a large body of empirical work that suggests the corrective safeguards have limited effectiveness. Finally, we analyze the advantages and disadvantages of the two corrective safeguards and outline directions for future empirical research.

*Keywords:* corrective safeguards, expert testimony, eyewitness evidence, eyewitness identifications, judicial instructions, jurors’ decision-making, legal precedents, legal safeguards.

*Index terms:* admissibility, expert testimony, eyewitness evidence, judicial instructions, legal safeguards, trial safeguards.
An Interdisciplinary and Cross-national Analysis of Legal Safeguards for Eyewitness Evidence

Mistaken eyewitness identifications can lead to wrongful convictions (e.g., Innocence Project, 2019; National Academies of Science [NAS], 2014), and it is now well-accepted that traditional trial safeguards, in particular, cross-examination, are ineffective at exposing the weaknesses of eyewitness identification (e.g., Devenport, Stinson, Cutler, & Kravitz, 2002). Thus, across the common law world, courts have increasingly turned to additional corrective trial safeguards to supplement the traditional reliance on cross-examination. At the same time, while some jurisdictions strengthened admissibility rules, these corrective safeguards are increasingly seen as an acceptable alternative to the exclusion of weak or unreliable identification evidence.

In this chapter, we conduct a cross-national comparison of legal safeguards used in Australia, Canada, and the United States of America. We briefly review the admissibility standards and then provide a more in-depth discussion of the implementation and effectiveness of two corrective safeguards—judicial instructions\(^1\) and expert testimony—that inform the jury of the factors that can make eyewitness identifications unreliable. We then evaluate these legal developments in light of contemporary psychological research on the effectiveness of these safeguards.

Wrongful Convictions, Eyewitness Evidence, and Jurors’ Decision-Making

Despite numerous exoneration cases, multiple inquiries (e.g., Devlin Committee, 1976; The Sophonow Inquiry, Cory, 2001), and extensive academic research, the evidence suggests

---

\(^1\)Terminology varies, however a distinction is sometimes drawn between judicial *directions* that a jury must follow (for example directions on the burden or standard of proof), and judicial *warnings or instructions* that inform, or warn the jury about matters that a jury can, or should, take into account when evaluating evidence that may be unreliable. We use “judicial instructions” to encompass any information the judge delivers about the eyewitness evidence to a jury. Thus, we use the term to include judicial warnings or directions (the terminology most commonly used in Australian jurisdictions) and judicial/jury instructions (the terminology more commonly used in the USA and Canadian decisions).
that criminal justice system personnel continue to underestimate the frequency of wrongful convictions as well as the role of mistaken eyewitness identification evidence as a contributing factor. Moreover, jurors have a limited understanding of factors that influence eyewitness reliability. Coupled with the belief that the knowledge of eyewitness factors is a matter of common sense, these preconceptions might influence the willingness of judges (and legislators) to utilize and to improve legal safeguards.

**Knowledge of the Frequency of Wrongful Convictions among Criminal Justice Officials**

Scholars have written about wrongful convictions for decades (e.g., Borchard, 1932); however, the introduction of the forensic DNA testing conclusively proved that factually innocent people had been wrongfully convicted in a significant number of cases. DNA testing uncovered numerous wrongful convictions in various countries and drew the public’s and officials’ attention to the problem (Ramsey & Frank, 2007). In the USA, at least 2,363 people have been exonerated since 1989 (National Registry of Exonerations, 2019), with 363 exonerated through DNA evidence (Innocence Project, 2019). In Canada, at least 22 people have been exonerated since 1993 (Innocence Canada, 2019). Although Australia does not have a definitive estimate, wrongful convictions have been established in more than 70 cases (71 cases between 1922 and 2015, Dioso-Villa, 2015; 84 cases since 1984, Hamer, 2019). Given that DNA exonerations are possible in a limited number of crimes and cases, some scholars argue that these numbers are just the tip of the iceberg and that the reality is much worse (see Hamer, 2014; Ramsey & Frank, 2007). Indeed, estimates of the prevalence of wrongful convictions vary from 0.5% to 20% of all felony cases in the USA (Ramsey & Frank, 2007) to 6% of all convicted people in England (Hamer, 2014). Clearly, the courts need to address the problem of wrongful convictions.
Knowledge of the problem is a necessary prerequisite to prevent miscarriages of justice; however, estimates of the prevalence of wrongful convictions by criminal justice personnel were positively biased toward their own jurisdiction and varied by their role. For example, when asked to estimate the percentage of all felony cases that resulted in a wrongful conviction in their jurisdictions (in Ohio, USA), the estimates ranged from 0.5% (prosecutors and police officers) to 1.0% to 3.0% (defense attorneys), with judges falling in the middle (0.5% to 1.0%, Ramsey & Frank, 2007). However, when asked about the frequency of wrongful convictions in the USA in general, all groups believed that they occur more often than in their own jurisdiction, with defense attorneys providing the highest estimates (4–5% of all felony cases). A similar pattern emerged in Michigan (USA): compared to defense attorneys, the police, prosecutors, and judges believed that errors within the criminal justice system occur less often (Smith, Zalman, & Kiger, 2011). Their interpretations of these errors varied even more drastically across their roles. Almost all defense attorneys surveyed (92%) stated that reforms of the criminal justice system were necessary to prevent wrongful convictions; however, only 10% of police officers, and 33% of prosecutors and judges reported that reform was necessary. The prosecutors and judges reported that, in their opinion, the majority of mistakes stem from negligence rather than intentional misconduct.

Nonetheless, analyses of factors contributing to wrongful convictions suggest that procedural issues, such as the misapplication of forensic science and mistaken eyewitness identifications, are more common than government misconduct (Innocence Project, 2019). Of particular relevance to this chapter, a substantial number of wrongful conviction cases included a mistaken identification by at least one eyewitness (30% of cases on the US National Registry of Exonerations, 2019; 72% of DNA exoneration cases, Innocence Project, 2019). Next, we explain
why eyewitnesses sometimes misidentify innocent people as the perpetrators and briefly discuss various factors that increase or decrease the reliability of eyewitness identifications. Basic knowledge of these eyewitness factors is necessary in order to understand the empirical evidence and to evaluate the effectiveness of the legal response to eyewitness evidence issues.

**Eyewitness Factors and their Effect on the Reliability of Identifications**

Eyewitness factors can be divided into two groups: estimator variables and system variables (Wells, 1978). Estimator variables include characteristics of an eyewitness (e.g., poor eyesight), of an event (e.g., presence of a weapon), and of a perpetrator (e.g., presence of disguise). These factors can be taken into account when evaluating eyewitness evidence but cannot be controlled by the criminal justice system because they occur at the time of the crime before the police are involved. On the other hand, system variables, such as who is in the lineup with the suspect, how the lineup is presented to the eyewitness, and what is said to the eyewitness before and after he or she views the lineup, are directly under the control of the criminal justice system. It is important to highlight that system variables can influence the eyewitness even when the lineup administrator did not intentionally bias the lineup. Whether intentional or not, suggestive identification procedures can affect not only the eyewitness’s immediate identification decision, but also her/his confidence in that decision (regardless of the actual accuracy), their judgments about testimony-relevant factors, and even their willingness to testify (Steblay et al., 2014; Wells & Quinlivan, 2009b).

Each estimator and system variable can either increase or decrease the reliability of eyewitness evidence. When examining how people (e.g., mock-jurors) evaluate eyewitness evidence, researchers typically manipulate the quality of the eyewitness evidence as a proxy for reliability. These manipulations relate to (a) the quality of the witnessing conditions (good vs.
poor), based on one or more estimator variables (e.g., exposure duration, distance, presence of a weapon), and/or (b) the quality of the identification conditions (non-suggestive vs. suggestive), based on one or more system variables (e.g., lineup presentation, lineup construction). Within studies, strong eyewitness evidence refers to good witnessing conditions and/or non-suggestive identification procedures, whereas weak evidence refers to poor witnessing conditions and/or suggestive identification procedures. Ideally, we want people to accept the evidence, or render guilty verdicts when the eyewitness evidence is strong (and, thus, more likely to be reliable), but not when the evidence is weak (and, thus, less likely to be reliable). See Table 1 for examples of witnessing and identification condition manipulations used in empirical studies investigating the effectiveness of legal safeguards. For interested readers, we also include a non-exhaustive list of key references from the eyewitness literature that demonstrate the effects of these manipulations on eyewitness decisions.

**Decision-Making in Cases that Involve Eyewitness Evidence**

In this section, we discuss jurors’ knowledge of eyewitness factors and examine whether their knowledge translates into an ability to determine the accuracy of an eyewitness. We also briefly explore factors that have been empirically shown to influence jurors’ assessments of eyewitness evidence.
Table 1

**Witnessing and Identification Conditions**

<table>
<thead>
<tr>
<th>Witnessing conditions</th>
<th>Good</th>
<th>Poor</th>
<th>Eyewitness research reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exposure duration</td>
<td>30 seconds or longer</td>
<td>Up to 12 seconds</td>
<td>Bornstein et al. (2012)*</td>
</tr>
<tr>
<td>Distance</td>
<td>Short distance</td>
<td>Long distance</td>
<td>Lindsay et al. (2008)</td>
</tr>
<tr>
<td>Presence of a weapon</td>
<td>No weapon</td>
<td>Weapon present</td>
<td>Fawcett et al. (2013)*</td>
</tr>
<tr>
<td>Crime seriousness</td>
<td>Serious crime</td>
<td>Crime not serious or eyewitnesses unaware of the seriousness</td>
<td>Leippe et al. (1978)</td>
</tr>
<tr>
<td>Stress</td>
<td>Low stress</td>
<td>High stress</td>
<td>Deffenbacher et al. (2004)*</td>
</tr>
<tr>
<td>Cross-race ID Disguise</td>
<td>No</td>
<td>Yes</td>
<td>Meissner &amp; Brigham (2001)*</td>
</tr>
<tr>
<td></td>
<td>Absent</td>
<td>Present</td>
<td>Morgan et al. (2013)</td>
</tr>
<tr>
<td>Identification conditions</td>
<td>Non-suggestive</td>
<td>Suggestive</td>
<td>Morgan et al. (2004)</td>
</tr>
<tr>
<td>Presentation of a suspect</td>
<td>Lineup</td>
<td>Showup</td>
<td>Steblay et al. (2003)*</td>
</tr>
<tr>
<td>Types of lineup</td>
<td>Sequential</td>
<td>Simultaneous</td>
<td>Steblay et al. (2011)*</td>
</tr>
<tr>
<td>Lineup construction</td>
<td>Suspect matched to the witness’s description and did not stand out</td>
<td>Suspect stands out (foil bias)</td>
<td>Clark et al. (2015)</td>
</tr>
<tr>
<td>Lineup instructions</td>
<td>Warning that the perpetrator may not be in the lineup</td>
<td>No warning or instruction bias</td>
<td>Steblay (1997)*</td>
</tr>
<tr>
<td>Lineup administration</td>
<td>Double-blind</td>
<td>Single-blind</td>
<td>Wells et al. (1998)</td>
</tr>
<tr>
<td>Post-identification feedback</td>
<td>No feedback</td>
<td>Confirmatory feedback</td>
<td>Kovera &amp; Evelo (2017)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Wells &amp; Bradfield (1998)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Charman &amp; Wells (2012)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Steblay et al. (2014)*</td>
</tr>
</tbody>
</table>

*Meta-analysis.
**Jurors’ knowledge of eyewitness factors.** Researchers have evaluated jurors’ knowledge of eyewitness factors using a number of investigative techniques, including surveys and experiments. One insightful way of examining people’s understanding is to compare potential jurors’ responses to those of experts with regard to a variety of questions or statements about eyewitness issues. Agreement between both groups would indicate that jurors share eyewitness experts’ opinions and, thus, have a good understanding of eyewitness issues. In contrast, disagreement would indicate that jurors lack awareness and/or an understanding of the relevant empirical evidence, which informed the experts’ responses.

In the first comparison of lay people and eyewitness experts, the two groups gave different responses for 15 out of 21 items (i.e., agreement for 29%, Kassin & Barndollar, 1992). For example, community members tended to overestimate the strength of the relationship between confidence and accuracy and underestimate the damaging influence of suggestive procedures (i.e., system variables), such as biased lineup instructions, biased lineup construction, and showups. Another study compared community members, judges, law enforcement personnel, and experts on knowledge of eyewitness factors (Benton, Ross, Bradshaw, Thomas, & Bradshaw, 2006). Community members and experts responded differently on 26 out of 30 items (agreement for 13%); troublingly, they disagreed on all eight system variable items. The judges and law enforcement personnel fared better—their responses differed from experts on 18 out of 30 items (agreement for 40%, Benton et al., 2006). These results mimic those found with a different methodology: when asked to list ten factors that they believed influenced eyewitness accuracy, the majority (94%) of participants’ responses related to estimator variables (Shaw, Garcia, & McClure, 1999).
Desmarais and Read (2011) portrayed a more optimistic picture in their meta-analysis of 23 published studies. The rate of agreement between lay people and experts was around 80% for factors such as the malleability of confidence, lineup instructions, wording of questions, and the influence of expectations on memory. However, the rate of agreement was lower for other crucial factors, such as the relationship between confidence and accuracy, cross-race identification, and the importance of fair lineup construction. Contrary to the previously-discussed studies, they concluded that lay people had better knowledge of system than estimator variables.

Nonetheless, knowledge of factors that influence eyewitnesses might not necessarily translate into better decisions when jurors have to assess various pieces of evidence and integrate this knowledge into their decisions. For example, jurors can correctly perceive a procedure to be more suggestive and still be more convinced by eyewitness evidence obtained from that procedure (Devenport et al., 2002). Taken together, the evidence suggests that people might have a limited understanding of how various factors affect eyewitnesses’ decisions, might underestimate the importance of system variables, and might not incorporate information about how identification procedures were conducted into their assessment of eyewitness evidence (Boyce, Beaudry, & Lindsay, 2007; Bornstein & Greene, 2017).

**Jurors’ sensitivity to eyewitness accuracy.** Rather than answer questions about eyewitness factors, the task of jurors in a real trial is much different. Ideally, jurors should be able to determine whether an eyewitness correctly identified the perpetrator (i.e., made a correct identification) or mistakenly identified an innocent person (i.e., made a false identification). To investigate how jurors evaluate correct and false identifications, researchers use a *genuine eyewitness paradigm*. In this paradigm, eyewitness-participants view a mock-crime, make an
identification, and then testify about the crime and their decision (Wells, Lindsay, & Ferguson, 1979). This approach has two clear benefits: (1) jurors are presented with real eyewitness decisions in the sense that the identification decision was genuine and not staged, and (2) the researchers know ground truth—whether the eyewitness made a correct or false identification. Thus, we can determine whether the jurors’ belief of the eyewitness was accurate (i.e., believed a correct identification) or inaccurate (i.e., believed a false identification).

A series of studies using the genuine eyewitness paradigm demonstrated people’s willingness to believe eyewitnesses. For instance, participants showed a general insensitivity to eyewitness accuracy with 80% believing the eyewitnesses regardless of their accuracy (Wells et al., 1979). Even after viewing a cross-examination of the witness, conducted by qualified lawyers or by law students, observers were not able to distinguish between correct and false identifications (Lindsay, Wells, & O’Connor, 1989). Overall, lay people consistently demonstrated insensitivity to eyewitness accuracy (Wells, Lindsay, & Tousignant, 1980; Lindsay, Wells, & Rumpel, 1981).

Semmler, Brewer, and Douglass (2012) proposed that jurors’ insensitivity to eyewitness accuracy might be an example of the fundamental attribution error (Ross, 1977; Chin & Crozier, 2018). Thus, one way to potentially improve jurors’ ability to evaluate eyewitness evidence might be to increase the salience of the situational factors (e.g., suggestive circumstances) that were present during the identification procedure. This can potentially be achieved by video-recording the identification procedure (Boyce et al., 2007; Wells & Wilford, 2013). Although there is promising evidence that viewing a video of an identification procedure can improve jurors’ ability to discriminate between correct and false identifications (Reardon & Fisher, 2011; Kaminski & Sporer, 2017), this sensitivity to eyewitness accuracy might be impaired if the video
captured suggestive procedures (Beaudry et al., 2015). Thus, it is critical that police use best practice identification procedures to minimize both deliberate and inadvertent suggestion.

**What influences jurors’ belief of eyewitnesses?** In order to improve jurors’ evaluation of eyewitness evidence, it is essential to understand which factors jurors take into consideration and whether these factors are good predictors of eyewitness accuracy. The most powerful predictor of jurors’ belief is the witness’s confidence at the time of the trial (e.g., Cutler, Penrod, & Dexter, 1990; Lindsay et al., 1981). No other variables consistently influence jurors’ decisions to the same extent as eyewitness confidence. Although an uncontaminated confidence statement obtained at the time of the identification can be a good predictor of accuracy (Brewer & Wells, 2006; Wixted & Wells, 2017), an eyewitness’s confidence during the trial is less informative (Bradfield & Wells, 2000; Brewer & Burke, 2002; for a review see Leippe & Eisenstadt, 2007) because factors such as preparing for the trial (Wells, Ferguson, & Lindsay, 1981) or receiving feedback about their identification (Wells & Bradfield, 1998) can inflate confidence without increasing accuracy.

Attempts to reduce jurors’ reliance on confidence, for the most part, have not been successful. For example, when an eyewitness was less confident at the time of an identification and more confident during the trial, observers ignored this confidence inflation unless it was explicitly questioned during cross-examination (Bradfield & McQuiston, 2004) or if the witness could not provide a compelling explanation (e.g., being nervous at the time of the identification) for the change in confidence (Jones, Williams, & Brewer, 2008). Instructing jurors not to use confidence also failed to have any effect (Fox & Walters, 1986). On the other hand, confidence inflation that was documented on the video (i.e., a video-recorded identification procedure in
which the eyewitness is not confident, as opposed to a confidence statement provided by the witness during the trial) decreased jurors’ belief of eyewitnesses (Douglass & Jones, 2013).

Towards Best Practice?

To this point, we have established that eyewitness evidence contributes to wrongful convictions, that police identification procedures can affect the reliability of eyewitness evidence, and that perceptions of this evidence do not necessarily align with cues to reliability. Crucially, police manuals across a number of jurisdictions do reflect—at least in part—the body of knowledge produced by eyewitness researchers. For example, in Australia, the New South Wales Police Force implemented non-suggestive instructions, double-blind lineup administration, and some components of the sequential lineup, requiring, wherever possible, to bring people into the witness’s view one by one rather than all at once (New South Wales Police Force, 2012; see Lindsay et al., 2009, for a discussion of the risks associated with using individual components rather than implementing the sequential “package”). Similarly, in Victoria, Australia, police members who take part in the investigation of the case are not allowed to conduct the lineup (Victoria Police Manual, 2003). In the USA, evidence-based practices have been recommended in numerous states and jurisdictions (see Innocence Project, 2019 for an up-to-date list).

The question remains, however, as to whether the regular practice follows the policy (Bertrand et al., 2018; Greene & Evelo, 2015). For this reason, it is essential to understand what happens when potentially unreliable evidence is presented before the jury. We focus our legal discussion on jury trials rather than judge-only trials for two reasons. First, admissibility rules and trial safeguards have risen in response to concerns about the jury. Second, the vast majority of empirical work into the evaluation of eyewitness evidence has examined mock-jurors rather
than judges (cf. Stinson, Devenport, Cutler, & Kravitz, 1997). We have demonstrated that jurors have a propensity to believe eyewitnesses regardless of the actual strength of their evidence (Boyce et al., 2007). Taken together, this creates risks in terms of ensuring a fair trial and the accuracy of verdicts. We next turn our discussion to a cross-national comparison of admissibility safeguards designed to prevent weak or unreliable eyewitness evidence from appearing in court and consider whether this safeguard has been an effective response to the dangers of mistaken eyewitness testimony.

**Admissibility Safeguards**

The risks associated with eyewitness evidence generally have not gone unnoticed; legal systems around the world have recognized that eyewitness identification evidence has a high chance of error. Courts have also acknowledged that eyewitness identification evidence raises particular difficulties for the fact finder given that it can be presented by a witness who is confident and honest, but mistaken (Australian Law Reform Commission [ALRC], 2005; Devlin Committee, 1976; NAS, 2014). Consequently, over time, courts and legislatures across jurisdictions have implemented a number of safeguards to manage both the admission and the evaluation of eyewitness identification evidence.

These legal safeguards fall into two broad categories. First, admissibility safeguards direct attention to whether the evidence should be *received* by the finder of fact (judge or jury) at all. In some jurisdictions, rules of admissibility will privilege certain types of identification procedures, such as identification parades, over others, or impose certain conditions, such as the expectation that the court will evaluate the reliability of the evidence prior to admission. At a more general level, and applicable to all evidence, common law jurisdictions also empower a trial judge to exclude evidence where the probative value of the evidence (i.e., the weight to be accorded to
This evidence) is outweighed by the risk of unfair prejudice to the accused, or where the admission of the evidence will give rise to an unfair trial (Edmond et al., 2013).

The second broad category of legal safeguards, which we have termed ‘corrective’ safeguards (the focus of the next section), includes judicial instructions and expert testimony. These safeguards are designed to assist the fact finder (usually the jury) to evaluate evidence after it has been admitted. In many cases, corrective safeguards are the preferred remedy in situations where the evidence has been admitted—often over objection—despite legitimate concerns about violations of rules and procedures or claims that the evidence may be unreliable. As we develop further below, this preference for relying on corrective safeguards over exclusion reveals a tension in how courts approach the risks associated with eyewitness evidence. On the one hand, Australian, Canadian, and American courts have acknowledged the importance of drawing on empirical research to inform the development and application of evidentiary rules and procedures. On the other, courts are reluctant to consider whether contemporary research supports claims that corrective safeguards are effective.

**Sufficiently Reliable? An Overview of Admissibility Safeguards**

In our comparison of admissibility in Australia, Canada, and the United States of America, we focus on the extent to which these courts have incorporated admissibility thresholds or tests that allow the judge to consider the reliability of the eyewitness evidence, and to then exclude unreliable evidence. We do not intend to provide a comprehensive and exhaustive account of admissibility safeguards, but to draw attention to a tendency to minimize the consequences of factors associated with reliability and instead rely on corrective safeguards to protect the defendant from risks associated with admitting evidence that might have low probative value.
Australia. Australia’s Uniform Evidence Law (UEL),\(^2\) initially drafted in 1985 and first enacted in 1995, adopted an approach to the admission of eyewitness testimony that discouraged in-court or dock identifications, and privileged the live identification parade, considered at the time to be the most reliable, and least prejudicial, form of eyewitness identification evidence (ALRC, 1985). At a general level, the Act prevents admitting evidence of eyewitness identification of the defendant (either in or out of court) unless that witness had participated in an identification parade including the defendant. However, of necessity, the UEL offers a number of exceptions to this requirement, including relieving investigators of the obligation to hold a parade in circumstances where the defendant refuses to cooperate, where it would be unfair to the defendant to hold a parade, or where it is not reasonably practical to hold a parade.\(^3\) Where one of these exceptions apply, investigators can turn to alternatives, most regularly using picture identification procedures (i.e., a photoboard/lineup; under section 115). At the State or Territory level, police guidelines or additional procedural legislation can govern the composition and conduct of identification procedures; however, a breach of these conditions will not necessarily result in exclusion (cf. the Crimes Act 1900 (ACT), section 233). Similarly, evidence of an identification should be excluded if it has been ‘intentionally influenced’; however, this has been interpreted narrowly, does not apply to picture identifications conducted under section 115, and

\(^2\) The Uniform Evidence Law regime governs the majority of Australian State and Territory jurisdictions, including the Commonwealth (Evidence Act 1995 (Cth)), New South Wales (Evidence Act 1995 (NSW)), Victoria (Evidence Act 2008 (Vic)), the Australian Capital Territory (Evidence Act 2011 (ACT)), the Northern Territory (National Uniform Legislation Act 2011 (NT)), and Tasmania (Evidence Act 2001 (Tas)). Note however that the Tasmanian Act has not incorporated the same provisions relating to identification evidence as the rest of the UEL jurisdictions. The States of South Australia, Queensland, and Western Australia remain governed by the common law, as modified by their own legislation. In this chapter we focus on the UEL jurisdictions, because together they cover the majority of criminal trials conducted in Australia.

\(^3\) Notwithstanding the emphasis on the live parade in the Act, very few, if any, live parades are now held in Australia. Indeed, when this requirement was proposed, it was acknowledged that the Australian Federal Police rarely held live parades (ALRC, 1985).
does not necessarily encompass live parades or photobords that were composed in a way that might have had a suggestive effect (e.g., where it was argued that the suspect stood out).

In all Australian jurisdictions, the identification evidence can also be challenged on the basis that the probative value is outweighed by the danger of unfair prejudice. For example, section 137 of the UEL could be used to exclude an in-court identification if a witness had rejected a parade or photoboard containing the defendant prior to trial. However, the effectiveness of this exclusionary rule has diminished with respect to an identification procured prior to the trial because Australia’s highest appellate court has now explicitly rejected reliability as a factor that judges can consider when determining the probative value of the evidence (IMM v The Queen, 2016). This has influenced the courts when considering identifications obtained in breach of police procedures or recognized good practice, in so far as such breaches are unlikely to result in the exclusion of evidence, even where those breaches mean that the evidence is of low probative value (e.g., The Queen v Dickman, 2017). Further, the UEL relies on a narrow definition of identification, meaning that a large amount of evidence relevant to the identification of the accused falls outside of the definition, and is subject to weak admissibility standards (San Roque, 2017). Overall, despite the apparently exclusionary orientation of the drive for reform behind the UEL, the trend in Australian decisions has been a preference to admit the eyewitness evidence and rely on corrective safeguards rather than exclude it.

**Canada.** Criminal law, evidence, and procedure in Canada are governed primarily by Federal (i.e., national) law. This can enhance consistency as compared to the United States, however, in contrast to Australia, Canada has not undertaken any significant legislative reform of the law of evidence. Consequently, Canada has not enacted any specific exclusionary provisions equivalent to the UEL, and the regulation of eyewitness identification remains governed
primarily by common law. For example, although section 6.1 of the *Canada Evidence Act* (1985), which applies in courts exercising Federal jurisdiction, allows witnesses to give evidence identifying the accused, it does not impose any conditions on such testimony.

Like Australia, Canadian decisions on admissibility (and as discussed below, on judicial instructions) incorporate references to eyewitness research. High-profile inquiries into wrongful convictions in Canada have recommended reforms to eyewitness procedures (e.g., the Sophonow Inquiry, Cory, 2001). Similarly, Canadian courts have developed a list of factors relevant to assessing the probative value of the evidence, including the risk of contamination or whether suggestive procedures were used (Hill, Tanovich, & Strezos, 2013). However, despite this expressed awareness, and in common with Australia, Canadian courts remain reluctant to exclude eyewitness evidence. And although individual jurisdictions or police forces have adopted procedural guidelines that govern the composition and administration of lineups, breaches of the guidelines in and of themselves do not necessitate exclusion (e.g., *R v Hibbert*, 2002; Campbell, 2018).

This reluctance could be because eyewitness identification is conceptualized far more explicitly as a form of permissible lay opinion, with the in-court identification still seen as the primary evidence, to be ‘corroborated’ by prior identification(s). The Canadian courts envisage that the witness will ordinarily express their opinion in the form of an in-court ‘dock identification’ (as contemplated in section 6.1) and that the probative value of this identification will be assessed by reference to the whole of the circumstances that have given rise to this identification (e.g., *R v Tat*, 1997). Ordinarily the in-court identification will be accompanied by the witness’s testimony that he or she has, on a previous occasion, identified the same person (e.g., during the police identification procedure); the fact finder can then determine the weight to
be given to the evidence, taking into account all of the circumstances, including those of any prior identification(s).

It is accepted that an in-court identification, standing alone, has very little probative value, and is insufficient grounds for a conviction (Hill et al., 2013); nonetheless, the Canadian approach to eyewitness identification presents as more permissive and inclusionary than that adopted in either Australia or the United States. For example, although Canadian case law contemplates the possibility that an in-court identification can be excluded, it permits witnesses to offer a positive in-court identification even if that witness had rejected a lineup containing the defendant during a prior formal identification procedure; Australian courts would be less likely to accept this evidence. Even highly suggestive or tainted identification evidence has been accepted as admissible (e.g., an identification made in court after the witness has viewed footage of the defendant in custody, R v Hibbert, 2002).

Canadian judges do have the capacity to exclude evidence where its probative value is outweighed by the danger of unfair prejudice to the accused (Hill et al., 2013) or direct an acquittal where it would be unreasonable for a jury to convict on eyewitness evidence alone (R v Hay, 2013). But, like Australia, an inclusionary orientation also leads to a preference for judicial instructions as the primary safeguard against the risk that a jury will place too much weight on the evidence. Consistent with this trend, the Canadian Charter of Rights and Freedoms (1982) opened up the possibility that the Charter might require, in some cases, that evidence be excluded where its admission might compromise the defendant’s right to a fair trial (Stuart, 2014). However, this remedy has not yet been extended to support a successful challenge to the admission of unreliable evidence, such as eyewitness identification evidence (e.g., Roach, 2007).
United States of America. Like Australia and Canada, the US jurisprudence has been influenced by judicial understandings of psychological research, but the multiplicity of jurisdictions in the US make for a more varied landscape. A distinctive feature of the US jurisprudence, however, is a more obvious preoccupation with the dangers associated with suggestive procedures. In line with the Supreme Court decision in *Manson v Braithwaite* (1977), if the defense believes that the identification procedure was suggestive or biased, the defense can file a motion-to-suppress evidence. In response, the judge will first determine whether the procedure was indeed unfair, and second whether the identification was nonetheless reliable, despite being obtained under suggestive circumstances.

Under the *Manson* criteria, the reliability of the witness is typically determined through an evaluation of five factors: opportunity to view the culprit, attention paid to the culprit, level of detail of the description of the culprit, the delay between the crime and the identification procedure, and the certainty of the witness. Not only are these criteria limited, but they also rely on an eyewitness’s self-report at trial and fail to appreciate that suggestive procedures themselves can corrupt that self-report (e.g., reporting higher confidence, a better view of the perpetrator, and longer exposure to the perpetrator, Wells & Quinlivan, 2009a). Although the emphasis in the US jurisprudence on preventing suggestive procedures is appropriate, it places far too great a reliance on judges’ knowledge of contemporary eyewitness research and ability to evaluate the evidence’s reliability. Further, in US courts, an acceptable remedy to a suggestive procedure is for investigators to hold a second ‘non-suggestive’ procedure, which is likely to compound, rather than reveal an error (Steblay & Dysart, 2016; Wells & Quinlivan, 2009b).

Recent decisions have, in part, recognized the weaknesses of the *Manson* criteria. For instance, in *New Jersey v Henderson* (2011) the court undertook a comprehensive review of
contemporary research and held that where the defense could point to (credible) evidence of suggestiveness, the prosecution needed to offer proof that the identification was, nonetheless, reliable. Similarly, *Oregon v Lawson* (2012) shifted the emphasis from judicial evaluation under the Manson criteria towards a more comprehensive pre-trial investigation that places the onus on the prosecutor to satisfy the court that all relevant criteria have been met. However, like *Henderson*, this ruling also places a significant burden on the defendant. Further, these rulings might not represent broader trends across US courts; notably, the Supreme Court in *Perry v New Hampshire* (2012) rejected arguments that the particular frailties of eyewitness identification gave rise to a general obligation to scrutinize the reliability of such evidence. Additionally, other jurisdictions have recently moved away from implementing much-needed reforms (Polimeni, 2018). Finally, like Australia and Canada, the Federal Rules of Evidence include a general discretion that permits the exclusion of evidence where the probative value is outweighed by the danger of unfair prejudice (Rule 403). This Rule, however, appears to be used sparingly to exclude eyewitness evidence that has either passed or bypassed an admissibility challenge (see related discussion in Tallent, 2011).

**A Preference for Corrective Safeguards?**

Troublingly, despite an acknowledged awareness of the role of eyewitness testimony in contributing to miscarriages of justice, none of the reviewed countries have developed a robust or consistent exclusionary framework to protect defendants from the dangers of unreliable eyewitness testimony. Although judges do provide reasons for their decisions, little is known about judges’ exclusion decisions in terms of patterns of decision-making more broadly. The only study, to our knowledge, that explored this topic demonstrated that, when evaluating a motion-to-suppress eyewitness evidence, US judges took into account the suggestiveness of how
the lineup was constructed and what instructions were given to the eyewitness before the identification procedure (Stinson et al., 1997).

Judicial approaches appear to be aligned with scholars who argue that excluding unreliable evidence deprives jurors of essential information. In their view, a properly instructed jury is expected to be able to evaluate the entirety of the evidence while being aware of the potential unreliability of admitted evidence (e.g., Woller, 2003). In all three countries, but perhaps especially in Australia and Canada, inclusionary decisions are justified on the basis that corrective safeguards are able to address the weaknesses in the evidence. This preference tends to ignore the interdependence of evidence—the fact that one piece of evidence can influence jurors’ evaluation of other pieces of evidence (Hasel & Kassin, 2009). And, as we turn now to a detailed examination of these corrective safeguards, we shall see that the judicial faith in the effects and efficacy of these safeguards—and, in particular, judicial instructions—is misplaced.

**Corrective Safeguards**

Two ‘corrective’ safeguards—judicial instructions and expert testimony—are increasingly being relied on as the preferred response to managing the risks associated with the admission of eyewitness identification evidence. This section provides an overview of the development of each of these corrective safeguards in Australia, Canada, and the USA. We then offer a general framework for assessing the effectiveness of these two corrective safeguards and turn to the empirical evidence examining the effectiveness of judicial instructions and expert testimony in shaping how jurors perceive eyewitness evidence and how the safeguards affect their verdicts.

Traditionally, common law courts have relied on the adversarial nature of the criminal trial, and in particular, cross-examination to reveal frailties or weaknesses in the evidence. In the case of eyewitness identification evidence, cross-examination might underscore inconsistencies and
demonstrate the unreliability of eyewitness evidence; however, it has been increasingly recognized that cross-examination is at best a partial safeguard (e.g., Edmond & San Roque, 2012; Righarts, O’Neill, Zajac, & Kovera, 2013).\(^4\) Courts have also acknowledged that cross-examination loses efficacy in situations where an eyewitness is giving honest and confident, but mistaken evidence. Thus, more recently, and in response to this awareness, courts have added the availability of cautionary instructions, directions, or warnings delivered by the judge to the fact finder (in this case, the jury). More recently still, expert testimony is used to assist the fact finder to understand the weaknesses specifically related to eyewitness identification. Both corrective safeguards provide general information about the psychology of eyewitness identifications and the factors that influence the reliability and accuracy of eyewitness evidence. We will discuss the features of both safeguards in greater detail below, but first it is essential to provide a framework for evaluating their effectiveness.

**Effects of Corrective Safeguards**

The key factor that should be kept in mind is that any educational aid is delivered to the jury. The jury consists of lay people, likely to have limited prior knowledge of the eyewitness literature, and each with their own background and beliefs. In examining the research on the effectiveness of any corrective safeguard, researchers and legal practitioners need to be attentive to the various effects that these safeguards can produce in a jury. Generally speaking, any corrective safeguard can lead to one of three outcomes: confusion, skepticism, or sensitivity (Cutler, Penrod, & Dexter, 1989). In order to be effective, a corrective safeguard should lead to a

---

\(^4\) This is putting aside the problem that an informed cross-examination relies on the lawyer knowing about the factors that influence the reliability of eyewitness evidence. As discussed above, knowledge of eyewitness factors may not be a matter of common sense (Boyce et al., 2007). Even when the cross-examination reveals factors that may undermine the reliability of eyewitness evidence, jurors may not incorporate this information into their evaluations because they are unaware of the importance of these factors (Devenport et al., 2002).
desirable effect, not just any effect; thus, unless a corrective safeguard increases jurors’ sensitivity, it should be considered ineffective.

**Confusion.** The term confusion is used to refer to jurors who assess the evidence in a way that contradicts information given in the corrective safeguard. That is, after hearing judicial instructions or expert testimony, jurors might render more guilty verdicts when the evidence is weak rather than strong. To illustrate, if jurors who heard judicial instructions believed an eyewitness who identified the perpetrator from a showup more often than an eyewitness who made an identification from a fairly-constructed lineup, then this would be labeled as confusion.

**Skepticism.** Skepticism is evident when jurors, after hearing a corrective safeguard, disbelieve an eyewitness regardless of the actual quality of witnessing and identification conditions or the accuracy of the identification. In other words, skepticism means that jurors reject both weak and strong eyewitness evidence.

**Sensitivity.** The desirable outcome when jurors are evaluating eyewitness evidence is sensitivity. If a corrective safeguard increases jurors’ sensitivity, it means that the jurors appropriately evaluated the quality of the eyewitness evidence and accepted strong, reliable evidence and dismissed weak, unreliable evidence.

Jurors can demonstrate sensitivity in two ways (Martire & Kemp, 2011): (a) sensitivity to witnessing and/or identification conditions, and (b) sensitivity to accuracy. Sensitivity to witnessing and/or identification conditions (WIC) refers to the jurors’ ability to correctly evaluate the situation under which the witness saw the perpetrator during the commission of the crime (i.e., witnessing conditions) and/or how the lineup procedure was conducted (i.e., identification conditions). If corrective safeguards are effective, the information from the expert or judge will increase jurors’ sensitivity to the witnessing and/or identification conditions; that is,
they will be more likely to believe eyewitnesses when the conditions were good (e.g., good lighting, long exposure to perpetrator, non-suggestive identification procedures) rather than poor (e.g., poor lighting, short exposure, suggestive identification conditions).

Even though, generally, eyewitnesses who had good witnessing conditions and made an identification under non-suggestive circumstances are more likely to be accurate, good witnessing and identification conditions should not be equated with eyewitness accuracy. Researchers typically examine sensitivity to witnessing and identification conditions using simulated eyewitness testimony (i.e., people acting as eyewitnesses and reading a script, or presenting participant-jurors with a transcript of a mock trial). There are two key limitations to simulated studies. First, they cannot replicate factors that are present in a real case, such as the eyewitness’s verbal and non-verbal behavior, which might provide additional cues to eyewitness accuracy (e.g., Kaminski & Sporer, 2018). Second, the concept of eyewitness accuracy itself does not exist in simulated studies because the eyewitness has been given a script and has not made an actual identification. As such, simulated testimony studies can speak to sensitivity to witnessing and identification conditions, but they cannot speak to mock-jurors’ sensitivity to eyewitness accuracy.

In real cases, we want jurors to believe accurate eyewitnesses and disbelieve inaccurate ones—sensitivity to eyewitness accuracy is the ideal outcome. Their decisions should be driven by the actual accuracy of the eyewitness rather than by the circumstances under which the witness saw and later identified the perpetrator. Given jurors’ general inability to determine eyewitness accuracy (e.g., Beaudry et al., 2015; Reardon & Fisher, 2011; Wells et al., 1979), effective corrective safeguards should increase jurors’ sensitivity to eyewitness accuracy.
The three potential outcomes in this framework—confusion, skepticism, and sensitivity—presume that corrective safeguards will have a significant effect on the jurors’ decisions. It is, of course, possible that the corrective effect will have no significant influence on jurors’ judgments; this is known as a null effect. Having established the potential effects of corrective safeguards on jurors’ evaluations of eyewitness evidence, we now outline the legal framework for judicial instructions and expert testimony in Australia, Canada, and the USA.

**Judicial Instructions**

Judicial instructions direct the jury to exercise caution in their evaluation of the eyewitness evidence. These instructions are generally delivered at the end of the trial, alongside directions on other matters (e.g., relevant law, the standard of proof). The content of judicial instructions about eyewitness evidence varies from one country to another (and within jurisdictions), and it is important to note that while there may be an obligation on the judge to provide the instruction, it remains a matter for the jury to determine what weight they ultimately give both to the evidence and to the factors detailed in the instruction.

**Australia.** Even prior to the enactment of the Uniform Evidence Law in 1995, Australian courts were ‘early adopters’ in terms of imposing a mandatory obligation on courts to warn juries about the frailties and risks associated with eyewitness identification evidence. The High Court emphasized that a failure to provide strong instructions to the jury could cause a miscarriage of justice and necessitate a new trial (*Domican v The Queen*, 1992). The UEL solidified this approach in sections 116 and 165 (and in Victoria in the *Jury Directions Act, 2015*), with section 116 framed as a mandatory obligation to direct a jury that ‘special caution’ is necessary when assessing eyewitness testimony.
The specific content of this mandatory instruction is shaped by guidance offered by Judicial Bench Books, which are regularly updated, and provide an extensive list of ‘standard’ or ‘suggested’ directions addressing a wide range of evidentiary matters that can be adapted to the circumstances of the case. For example, the New South Wales [NSW] direction (implementing section 116) includes a generalized warning that mistaken eyewitness evidence has been implicated in wrongful convictions, and that a witness may be sincere, honest, and impressive yet still be mistaken (Judicial Commission of NSW, 2007). Beyond this, however, the judicial instructions suggested in the NSW Bench Book tend to focus on estimator variables, perhaps on the assumption that risks associated with system variables (such as suggestive procedures) will be excluded from the court through the application of admissibility rules. Similarly, the judicial instructions used in Victoria (Judicial College of Victoria, Australia, 2006) warn jurors that identification evidence is potentially unreliable and then ask them to consider a variety of eyewitness factors grouped into three sections: circumstances of observation, factors concerning the witness, and factors concerning the identification. Each section poses several questions to the jury; for example: “How far away was the witness from what s/he was observing?” or “Was the identification process conducted fairly? For example, did the other people in the photoboard look sufficiently similar to the accused?” However, they do not provide any explanations for why those factors are important, sometimes provide confusing guidelines for the evaluation of the evidence, and do not always represent the research accurately. For instance, the Victorian example instructs the jury to consider: “Was the witness stressed or fearful at the time of the observation? If so, what effect would this stress have had on him/her? For some people, their powers of observation increase under stress. Others black out and their powers of observation diminish. You need to decide how the witness is likely to have reacted in this case.” In contrast
to these instructions, the research has shown that stress decreases identification accuracy and makes an eyewitness more prone to misinformation (Deffenbacher, Bornstein, Penrod, & McGorty, 2004; Morgan et al., 2004; Morgan, Southwick, Steffian, Hazlett, & Loftus, 2013).

Finally, it is important to note that even when judicial instructions are very detailed and reveal multiple weaknesses in the evidence, it remains open to the jury to nonetheless accept the evidence as probative of the accused’s guilt (see e.g., Dickman, 2017). It is only in rare cases, such as where the identification evidence is the sole evidence against the accused, that the jury needs to be convinced beyond reasonable doubt of the accuracy of any particular identification (rather than the totality of the evidence) to render a guilty verdict.

Canada. Like Australia, Canada has tended to prefer the use of judicial instructions over exclusion. The leading Supreme Court case of *R v Hibbert* (2002) was primarily about the sufficiency of the directions given by the trial judge, rather than a case about the exclusion of the evidence. In *Hibbert*, the evidence of multiple witnesses had been compromised by pre-trial exposure to images of the defendant in the media. In addition, a number of witnesses had rejected a lineup containing the defendant in formal identification procedures but were nonetheless permitted to identify the defendant in court. Notwithstanding these serious problems, the Supreme Court decision was whether the trial judge had explained adequately to the jury the weaknesses in the evidence, the risks of displacement (also known as mugshot exposure, e.g. Deffenbacher, Bornstein, & Penrod, 2006), and the consequent need for the jury to exercise special caution when evaluating the evidence. The judgments in *Hibbert* offer a detailed discussion of the factors that the fact finder ought to take into account when assessing the reliability of the evidence, many of which are based on eyewitness research into estimator and system variables. The decision also recognizes that there is a lack of correlation between
confidence and accuracy and that an in-court identification, standing alone, is of almost negligible probative value.

These judicially-recognized factors developed via the case law have been incorporated into the Canadian Model Jury Instructions, which (like the examples from Australia discussed above) include standard or suggested instructions that refer to the need for special caution, and the association between mistaken eyewitness identification and wrongful convictions (Canadian Judicial Council, 2012). Thus, the Canadian instructions provide a list of factors that should be taken into account depending on the circumstances of each case. These factors include witnessing conditions (e.g., how long the witness observed the offender, lighting, distance), as well as whether there was a risk of contamination (e.g., _R v McIntosh, 1997_), or whether the identification procedures might have been suggestive. However, as discussed above, the Canadian instructions provide little guidance to the jury regarding the application of the instructions and tend to present the impact of the various variables as matters of common knowledge. There is little or no commentary provided in the Model Directions to assist the judge in framing the significance of each factor, and minimal reference to the relevant literature or contemporary research. However, in contrast to the Australian guides, the Canadian Model Directions draw specific attention to the risks of displacement and other system variables (such as the fairness of the array). However, these inclusions are double-edged because they point to the overtly inclusionary nature of the Canadian admissibility standards. Furthermore, the details of the instructions remain up to the individual judge’s discretion, monitored by the appeal courts on an ad-hoc basis. This approach is problematic considering judges have a limited understanding of eyewitness issues (Benton et al., 2006).
United States of America. The USA’s *Telfaire* judicial instructions are arguably the most widely discussed and most commonly researched judicial instructions in the literature (*US v Telfaire*, 1972). These instructions ask jurors to consider factors such as the duration of the event, lighting conditions, familiarity with the perpetrator, delay, and the credibility of the witness (i.e., estimator variables). They also briefly mention identification conditions: “If the identification by the witness may have been influenced by the circumstances under which the defendant was presented to him for identification, you should scrutinize the identification with great care” (Greene, 1988, p. 3). The *Telfaire* instructions have been widely criticized for three main reasons (e.g., Dufraimont, 2008; Sheehan, 2011). First, they are based on previous cases and, thus, often do not reflect findings of psychological research (Cutler & Penrod, 1995). Second, even though the instructions mention a number of important factors, they do not explain how these factors might affect an eyewitness. Third, they do not warn jurors that even those eyewitnesses who appear highly confident at the time of the trial might be mistaken.

Taking into account these critiques, the New Jersey Supreme Court developed its own set of instructions—the *Henderson* instructions (*New Jersey v Henderson*, 2011). These instructions were informed by the Special Master’s extensive review of eyewitness research and come closest to accurately reflecting findings of eyewitness research. The *Henderson* instructions provide a comprehensive overview of eyewitness factors as well as an explanation of how each of the factors can influence an identification. It is important to note that not all jurisdictions in the USA have even accepted that judicial instructions form a necessary part of a trial that has relied on eyewitness identification evidence. Thus, one of the NAS’s key recommendations was a requirement that model, national judicial instructions be formulated and mandated (NAS, 2014).
What does the empirical evidence say about the effectiveness of judicial instructions?

The majority of empirical studies published in peer-reviewed journals evaluated the effectiveness of judicial instructions used in the USA; only one study to date tested Australian (New South Wales) judicial instructions. Overall, the empirical evidence suggests that judicial instructions regarding eyewitness identification evidence produce either skepticism or have little effect on jurors’ judgments (see Table 2). We highlight a few key studies here to illustrate how researchers examined this issue and to give us an opportunity to discuss more detailed results.

Greene (1988) tested the effectiveness of the Telfaire judicial instructions (vs. control) using a videotaped trial in which she manipulated the witnessing conditions. Rather than increasing mock-jurors’ sensitivity to the quality of the witnessing conditions (i.e., greater belief for good vs. poor conditions), the Telfaire instructions led to skepticism (i.e., reduced belief regardless of witnessing conditions). In an attempt to improve the judicial instructions, in a second study, Greene (1988) developed a revised and simplified set of Telfaire instructions. In addition, Greene added information about factors such as: the influence of stress, lineup fairness, and the relationship between accuracy and confidence. Jurors who heard the revised instructions demonstrated better knowledge of eyewitness factors and rendered fewer guilty verdicts (cf. the Telfaire instructions). Again, however, jurors were not sensitive to the quality of witnessing conditions and showed skepticism towards the evidence. It is worth noting that regardless of the presence of any instructions, none of the juries returned a guilty verdict when the evidence was weak and only 24% of the juries convicted the defendant in the strong version.
<table>
<thead>
<tr>
<th>Author</th>
<th>Safeguard</th>
<th>Evidence</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Greene (1988); study 1</td>
<td>None vs. <em>Telfaire</em></td>
<td>Evidence: weak vs. strong (i.e., lighting, distance, clarity of view)</td>
<td>Skepticism</td>
</tr>
<tr>
<td>Greene (1988); study 2</td>
<td>None vs. <em>Telfaire</em> vs. revised and simplified instructions</td>
<td>Same as study 1</td>
<td>Skepticism (Telfaire and revised instructions)</td>
</tr>
<tr>
<td>Cutler et al. (1990)</td>
<td>None vs. expert testimony vs. <em>Telfaire</em> instructions</td>
<td>WIC*: poor vs. good (i.e., presence of disguise, weapon presence, length of retention interval, lineup instructions)</td>
<td>Skepticism (expert testimony) Null effect (judicial instructions)</td>
</tr>
<tr>
<td>Ramirez et al. (1996); study 1</td>
<td>None vs. <em>Telfaire</em> instructions. Safeguard timing: before and after evidence vs. before vs. after</td>
<td>WIC*: poor vs. good (i.e., disguise, weapon presence, delay: two days or two weeks, lineup presentation: biased vs. unbiased)</td>
<td>Skepticism if the judicial instructions are presented after the evidence or overbelief when presented before and after the evidence</td>
</tr>
<tr>
<td>Ramirez et al. (1996); study 2</td>
<td>Safeguard type: none vs. <em>Telfaire</em> instructions vs. revised <em>Telfaire</em> instructions.</td>
<td>WIC*: poor vs. good (i.e., exposure duration, distance, stress, weapon presence, lighting, delay: two hours vs. two weeks, mugshot exposure, lineup instructions, lineup construction)</td>
<td>Skepticism (Telfaire) Null effect (revised instructions)</td>
</tr>
<tr>
<td>Martire &amp; Kemp (2009)</td>
<td>None vs. congruent expert testimony vs. incongruent expert testimony vs. judicial instructions</td>
<td>Eyewitness accuracy: accurate vs. inaccurate</td>
<td>Null effect (all manipulations)</td>
</tr>
<tr>
<td>Papailiou et al. (2015)</td>
<td><em>Henderson</em> instructions vs. minimal eyewitness instructions</td>
<td>Identification quality: weak vs. strong (operationalized by 10 parameters)</td>
<td>Skepticism (<em>Henderson</em>)</td>
</tr>
<tr>
<td>Dillon et al. (2017)</td>
<td>None vs. <em>Henderson</em> instructions before eyewitness testimony vs. <em>Henderson</em> instructions after eyewitness testimony vs. control</td>
<td>Witnessing conditions: poor vs. good (i.e., delay, weapon presence, exposure duration). Identification conditions: poor vs. good (lineup type, lineup instructions, presence of feedback)</td>
<td>Skepticism (all manipulations)</td>
</tr>
<tr>
<td>Jones et al. (2017)</td>
<td>None vs. expert testimony vs. <em>Henderson</em> instructions vs. enhanced instructions vs. expert testimony and <em>Henderson</em> instructions</td>
<td>Witnessing conditions: poor vs. good (i.e., exposure duration, weapon presence, delay) Identification conditions: poor vs. good (lineup type, lineup instructions, presence of feedback)</td>
<td>Skepticism (expert testimony) Null effect (both types of judicial instructions)</td>
</tr>
</tbody>
</table>

*Witnessing and identification conditions.*
Cutler, Dexter, and Penrod (1990) also tested the effectiveness of the *Telfaire* judicial instructions and of expert testimony (the results of expert testimony will be reported in the corresponding section), compared to a control condition. They manipulated the quality of witnessing and identification conditions and found that jurors were insensitive to the quality of eyewitness evidence, regardless of the presence of the instructions.

Ramirez, Zemba, and Geiselman (1996) compared the *Telfaire* instructions with revised *Telfaire* instructions (simplified and offered guidance on how jurors should evaluate eyewitness factors) compared to a control condition. They also manipulated the quality of witnessing and identification conditions. Without the instructions, jurors were sensitive to the quality of eyewitness evidence—they rendered more guilty verdicts when the witnessing conditions were good rather than poor. Contrary to this desired outcome, the *Telfaire* instructions resulted in skepticism. The revised instructions, on the other hand, preserved jurors’ sensitivity and improved their knowledge of eyewitness factors.

The *Telfaire* instructions might be less effective than anticipated because of the content and/or its presentation (i.e., as a list with little elaboration). Theoretically, the newer *Henderson* instructions should be more effective in sensitizing jurors to the quality of eyewitness evidence because they provide an empirically-based description of eyewitness factors. Nonetheless, several tests of the *Henderson* instructions found evidence of skepticism: after hearing these instructions, jurors disbelieved the eyewitnesses, regardless of the quality of witnessing and identification conditions (Dillon, Jones, Bergold, Hui, & Penrod, 2017; Papailiou, Yokum, & Robertson, 2015).

Jones, Bergold, Dillon, and Penrod (2017) revised the *Henderson* instructions to explain how every factor might influence eyewitness accuracy by explicitly referring to research findings
(e.g., “A meta-analysis on mug shot effects found that seeing someone in a mug shot can double the risk of a subsequent misidentification. Specifically, false identifications were 16% for a previously unseen face, but went up to 38% when witnesses previously viewed a mug shot of that person,” Jones, 2016, p. 100). Their study investigated whether general Henderson instructions, their enhanced Henderson instructions, or the combination of Henderson instructions and expert testimony improved mock-jurors’ evaluation of eyewitness evidence compared to the no corrective safeguard control condition. Mock-jurors heard the instructions before the eyewitness testimony and also received a printed copy of the instructions.

Despite being evidence-based, compared to the control condition, none of the variations of the instructions affected jurors’ verdicts nor did they improve jurors’ knowledge of the majority of eyewitness factors present in the study. Regardless of the instruction condition, the quality of witnessing conditions did not affect jurors’ judgments of defendant guilt. Contrary to Jones et al.’s expectations, jurors were sensitive to the quality of identification conditions without any corrective safeguards, rendering more guilty verdicts when the identification conditions were good rather than poor (Jones et al., 2017). Taken together, there is minimal empirical evidence that the Henderson instructions improve jurors’ decisions.

Thus far, the studies we discussed simulated eyewitness evidence by presenting scripted eyewitness testimonies to mock-jurors. This type of research examines whether judicial instructions help jurors evaluate the quality of witnessing and/or identification conditions; however, it does not answer the ultimate question of whether judicial instructions improve jurors’ sensitivity to eyewitness accuracy. The only study to date that speaks to this question tested the judicial instructions used in New South Wales, Australia (Judicial Commission of NSW, 2006). In this study, Martire and Kemp (2009) presented video-recorded testimony of
genuine eyewitnesses to examine sensitivity to identification accuracy, but did not manipulate
the witnessing or identification conditions. Mock-jurors were sensitive to identification
accuracy—they were correct 64% of the time—but the judicial instructions did not improve their
sensitivity.

Overall, to date, there is no support for the long-held belief that judicial instructions are an
effective corrective safeguard for eyewitness evidence. There is no evidence that judicial
instructions improve jurors’ sensitivity to either the accuracy of the eyewitness identification or
the relevant factors (i.e., witnessing and identification conditions) that affect the reliability of
eyewitness evidence. One limitation of this body of work is that the research has examined only
a few of the available instructions (e.g., Telfaire, Judicial Commission of NSW, Henderson),
while neglecting others (e.g., judicial instructions used in Canada or other Australian
jurisdictions). However, considering that the Henderson instructions, which incorporate
psychological research and provide explanations for each of the eyewitness factors, were largely
ineffective, it is unlikely that any other judicial instructions, presented in the same way, would be
more effective.

**Expert Testimony**

Expert testimony about the psychology of eyewitness identification is another available
corrective safeguard that can be used alongside judicial instructions. An expert’s testimony will
usually include a brief overview of the way human memory works and a discussion of factors
that influence the reliability and accuracy of eyewitness evidence. Rather than providing a list of
factors, experts explain why certain factors are important and sometimes present results of the
studies that informed their claims. In contrast to judicial instructions, experts do not read this
information to the jury; but rather provide it in response to questioning by one of the attorneys.
Also in contrast to judicial instructions, expert testimony is subject to cross-examination; any claims made by the expert can be disputed by the opposing side.

As a matter of general principle, witnesses are confined to giving evidence of their observations and of facts, and are prohibited from giving evidence in the form of an opinion (e.g., UEL, section 76). However, all three countries under consideration provide for exceptions, including an exception that allows for the court and fact finder to receive expert opinion evidence. Broadly speaking, an expert is a witness who can demonstrate that he or she possesses some form of specialized knowledge based on training, study or experience; and the opinion she or he proffers is based on that specialized knowledge (see e.g., UEL, section 79; Federal Rules of Evidence, rule 702). While controversies have arisen in many jurisdictions about the reliability and validity of expert opinion evidence in specific domains (see for e.g., National Research Council, 2009), it is well established that parties can call appropriately qualified expert witnesses to prove matters relevant directly to a fact in issue (e.g., a fingerprint expert). What is less accepted is whether the expert can be allowed to offer information to a jury relevant to the assessment of the credibility or reliability of an eyewitness. Overall, courts have been historically reluctant to admit evidence from experts in these circumstances preferring instead to rely on judicial instructions. In general, this attitude was informed by the (we would argue, erroneous) belief that many of these matters could be adequately explored in cross-examination and/or were simply matters of general knowledge (for a discussion see Boyce et al., 2007; Bornstein & Greene, 2017).

Historically, Australia, Canada, and the USA have approached the question of admissibility of expert evidence with general reluctance and a desire to preserve the decision-making role of the jury. However, more recently, the three countries have diverged in their
willingness to allow experts to testify. In the USA, despite the lack of national conformity, the inclusion of expert testimony that touches on the credibility or reliability of eyewitness evidence is not generally regarded with the same hostility as it is in Canada. Australia falls somewhere in-between, with a growing acceptance that in some cases it is appropriate to call an expert to provide general information about eyewitness or memory issues to assist the jury or judge to evaluate evidence. However, it is important to note that even when experts are admitted, they are limited to presenting general information, and cannot be called on to proffer an opinion as to whether particular factors have definitively affected the reliability of the particular evidence, or whether a particular witness can be considered reliable.

**Australia.** In UEL jurisdictions, expert evidence about eyewitness identification can be admitted either under section 79 (the exception to the opinion rule allowing experts to proffer opinion evidence) or via section 108C (a specific exception to the general rule that prohibits the admission of evidence relevant only to the credibility of a witness). Section 108C incorporates the same criteria for the expert as those outlined in section 79, but with the additional criteria that the evidence must have the capacity to “substantially affect the assessment of the credibility of the witness.” (*Evidence Act 1995 (NSW)*). Although the introduction of this section (and its amendment in 2009) was primarily driven by concerns about the evaluation of witnesses in child sexual assault cases, it is not limited to these cases. Coupled with the abolition of the ‘common knowledge rule’ in section 80 of the UEL, the introduction of section 108C indicates a liberalization of approach, as compared to the traditional position. Consequently, Australian courts no longer hold that eyewitness factors are matters of common knowledge and are now amenable to expert evidence (Freckelton, 2014).
Although uptake of the opportunity offered by the insertion of section 108C has been slow (New South Wales Law Reform Commission [NSWLRC], 2012), expert evidence, proffered by a psychologist with expertise in eyewitness testimony and memory has been admitted in a number of Australian cases, across jurisdictions, including in *Bayley v R* (2016), *Gittany v R* (2016), and *Dupas v R* (2012). In particular, the admission of expert evidence in eyewitness cases has been influenced by the growing jurisprudence (in cases involving child sexual assault) that accepts that—in some cases—there is a need to offer juries information to counter what is considered to be popular misconceptions about the credibility of a witness and/or the behavior of a witness after experiencing a criminal event. Thus, for example, expert evidence can be targeted towards ensuring that juries properly understand the relationship between confidence and accuracy and countering the common lay misconception that confidence at the time of the testimony can be correlated in a simplistic way with accuracy (see e.g., Brewer & Wells, 2006). However, as noted above, sections 79 and 108C limit the scope and content of evidence to relatively general information about the state of the research rather than on factors that may be seen to be present in the particular case, leaving the jury (or judge) to consider the impact of those factors (*Dupas*, 2012; *Gittany*, 2016).

**Canada.** In contrast to Australia (and to some extent the United States), Canada has remained hostile to the admission of expert testimony on eyewitness evidence. Canadian courts have accepted psychologists as expert witnesses in relation to substantive issues, such as the assessment of the mental state or condition of the accused at the time of the defense, but as Chin and Crozier (2018) argue, have tended to prefer ‘dispositional’ evidence over evidence that illuminates the psychology of decision-making more broadly. In cases involving the admission of expert testimony, Canadian courts have held more strongly than the other countries to the
position that information about the psychology of eyewitness identification is within the common knowledge (or ‘ken’) of the decision-maker. Overall, courts have expressed a strong preference for such information to be delivered by the judge presiding over the trial, rather than allowing a party (usually the defense) to call an eyewitness expert.

The influential case of R v McIntosh (1997) exemplifies this conservative approach. The judge in that case not only rejected the expert on the grounds that their evidence did not meet the criteria for the admission of expert evidence established in R v Mohan (1994), but also emphasized that the information that the expert would provide was not “outside of the normal experience of the trier of fact.” Although the Sophonow Inquiry (conducted in 2000), chaired by Justice Cory in the wake of a miscarriage of justice in which eyewitness testimony was seriously implicated, did recommend that courts “readily admit properly qualified expert evidence pertaining to eyewitness identification” (Cory, 2001, p. 18) to assist the fact finder, this was not taken up in the subsequent Report on the Prevention of Miscarriages of Justice (Department of Justice, 2004). Both this report and the subsequent review (Department of Justice, 2011), confirmed the preference for judicial instructions over the admission of expert testimony to assist the fact finder. In addition to expressing the view that the widespread knowledge of the frailties of eyewitness identification (amongst both lay people and the judiciary) rendered expert evidence ‘redundant,’ the report also reinforced the traditional position that allowing an expert to testify about the potential frailties of a particular witness’s testimony risked usurping role of the jury as the fact finder (Department of Justice, 2004). Thus, the concern remains that the jury will effectively defer to the expert’s account of the risks associated with this type of evidence, without undertaking their own evaluation of the specific witness. Although more recent cases in Canada have adopted a more inclusionary approach (e.g., Campbell, 2018) and have permitted
an expert to proffer an opinion, these cases appear to be outliers and cannot yet be taken as
evidence of a trend towards a more general acceptance of expert evidence.

**United States of America.** Earlier cases in the USA tended to place greater reliance on
the adversarial safeguard of defense cross-examination as the primary mechanism through which
the reliability of evidence of incriminating eyewitness testimony would be challenged and its
weaknesses revealed to the fact finder. However, more recent cases have recognized that this will
not always be sufficient to articulate to the jury the real need for caution. Thus, in contrast to
Canada, a number of jurisdictions have become, in recent years, far more open than in the past to
the admission of expert testimony to guide the jury in its assessment of eyewitness testimony.

This shift in allowing expert testimony corresponds with the growing acceptance of the
need to ensure that the jury is properly instructed about the frailties of eyewitness testimony.
This position is not uniform, but a number of key recent decisions in State courts have rejected
the proposition that expert evidence is never admissible, and have instead held that the trial judge
has the discretion to admit such evidence from a properly qualified expert, provided that the
evidence does not extend to proffering a direct opinion as to the reliability or credibility of the
specific witness and their testimony (e.g., *Commonwealth v Walker*, 2014; Polimeni, 2018). And
in an outlier decision, a Federal Circuit court has even permitted an expert to conduct an
experiment using the evidence presented at trial, and proffer an opinion tailored to that evidence
(*Newsome v McCabe*, 2003). More conventionally, influential cases such as *New Jersey v
Henderson* (discussed above), advocated a greater use of expert testimony, as has the NAS
(2014; see also Tallent, 2011).

**What does the empirical evidence say about the effectiveness of expert testimony?**
Investigations into the effectiveness of expert testimony in sensitizing jurors to witnessing and
identification conditions provide mixed conclusions (see Table 3 for a summary). On the one hand, some studies found that expert testimony produced skepticism. That is, compared to jurors in the control condition, those who heard expert testimony gave lower ratings of guilt and deemed the eyewitness to be less accurate, regardless of the quality of witnessing and/or identification conditions (Cutler, Dexter, & Penrod, 1990; Jones et al., 2017), or eyewitness confidence (Fox & Walters, 1986). On the other hand, some studies found that expert testimony improved sensitivity to witnessing and/or identification conditions (Cutler, Penrod, & Dexter, 1989; Geiselman & Mendez, 2005). Another study found more complicated results, with jurors in the expert testimony (cf. control) condition demonstrating an increased awareness of some suggestive procedures (i.e., instruction, but not foil, bias); however, this awareness did not lead to a significant difference in the jurors’ verdicts (Devenport et al., 2002).

Furthermore, a number of studies have revealed both skepticism and sensitivity. For example, specific expert testimony that commented on factors present in the case successfully sensitized jurors to witnessing conditions (Geiselman et al., 2002, Experiment 1). However, a more traditional general expert testimony—in line with the accepted approach in Australia, Canada, and the USA—did not influence jurors’ judgments. Geiselman et al. suggested that the attorneys’ closing arguments directed jurors’ attention to missing pieces of evidence rather than to the quality of eyewitness evidence. To add to the mixed results, yet other studies found that expert testimony led to confusion (i.e., jurors evaluated the evidence contrary to the information conveyed by the expert; Lindsay, 1994, Experiment 5), or had a null effect on jurors’ judgments (Devenport & Cutler, 2004).

Taken together, the results of expert testimony studies do not warrant a definitive conclusion regarding the effectiveness of expert testimony in sensitizing jurors to the quality of
witnessing and identification conditions. These mixed results might be caused by differences in expert testimony content and presentation. Even though experts explain the psychology of eyewitness identifications, individual differences between experts would mean that the delivery of information will vary, potentially making expert testimony effective in some instances, but not in others.

More studies have looked at jurors’ sensitivity to witnessing and identification conditions, but few studies have examined whether expert testimony improves jurors’ sensitivity to eyewitness accuracy. The first study of this kind presented mock-jurors with expert testimony and the cross-examination of genuine eyewitnesses, who viewed the crime under poor, moderate, or good witnessing conditions (Wells et al., 1980). Expert testimony did not sensitize them to witnessing conditions, instead it produced skepticism by decreasing ratings of believability even for accurate eyewitnesses. Expert testimony did, however, reduce the correlation between mock-jurors’ willingness to believe the eyewitness and the eyewitness’ confidence reported during cross-examination. This reduced reliance on an eyewitness’s confidence at the time of the testimony is an important improvement, but it we should highlight that an eyewitness’s uncontaminated confidence statement at the time of the identification does relate to accuracy (e.g., Wixted & Wells, 2017). Martire and Kemp’s (2009) study also demonstrated that expert testimony did not sensitize mock-jurors to eyewitness accuracy.
Table 3

Summary of Studies that Tested the Effectiveness of Expert Testimony

<table>
<thead>
<tr>
<th>Author</th>
<th>Safeguard</th>
<th>Evidence</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wells et al. (1980)</td>
<td>None vs. expert testimony</td>
<td>Eyewitness accuracy: accurate vs. inaccurate</td>
<td>Skepticism</td>
</tr>
<tr>
<td>Fox &amp; Walters (1986)</td>
<td>None vs. general expert testimony vs. specific expert testimony</td>
<td>Eyewitness confidence: high vs. low</td>
<td>Skepticism (general and specific expert testimony)</td>
</tr>
<tr>
<td>Cutler, Dexter, &amp; Penrod (1989)</td>
<td>None vs. expert testimony. Eyewitness confidence: 80% vs. 100%</td>
<td>WIC*: poor vs. good (i.e., presence of disguise, weapon presence, length of retention interval, lineup instructions)</td>
<td>Sensitivity (strength of the prosecution’s and of the defense’s cases, did not affect verdict)</td>
</tr>
<tr>
<td>Cutler, Penrod, &amp; Dexter (1989)</td>
<td>Expert testimony: descriptive vs. quantified. Expert opinion: expert opinion vs. no opinion.</td>
<td>WIC*: poor vs. good (i.e., presence of disguise, weapon presence, length of retention interval, lineup instructions) Witness confidence: 80% vs. 100%</td>
<td>Sensitivity (verdict and belief, both types of expert testimony)</td>
</tr>
<tr>
<td>Cutler et al. (1990)</td>
<td>None vs. expert testimony vs. Telfaire instructions.</td>
<td>WIC*: poor vs. good (i.e., presence of disguise, weapon presence, length of retention interval, lineup instructions) Witness confidence: 80% vs. 100%</td>
<td>Skepticism (expert testimony) Null effect (judicial instructions)</td>
</tr>
<tr>
<td>Devenport et al. (2002)</td>
<td>None vs. expert testimony</td>
<td>Identification conditions: foil bias, lineup instructions, lineup presentation</td>
<td>Sensitivity (lineup instructions, did not affect verdict)</td>
</tr>
<tr>
<td>Geiselman et al. (2002, Experiment 1)</td>
<td>None vs. general expert testimony vs. specific expert testimony</td>
<td>WIC*: poor vs. good (i.e., lighting, distance, duration of view, emotional state of the witness, cross-race identification, presence of disguise, suggestiveness of lineup instructions, showup or lineup, same day or two-day delayed identification)</td>
<td>Sensitivity (specific expert testimony) Null effect (general expert testimony)</td>
</tr>
<tr>
<td>Author</td>
<td>Safeguard</td>
<td>Evidence</td>
<td>Findings</td>
</tr>
<tr>
<td>-------------------------</td>
<td>----------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>Devenport &amp; Cutler (2004)</td>
<td>None vs. defense-only expert vs. opposing experts</td>
<td>Foil bias: biased vs. unbiased</td>
<td>Null effect</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lineup instructions: suggestive vs. non-suggestive</td>
<td></td>
</tr>
<tr>
<td>Leippe et al. (2004)</td>
<td>Expert testimony pre-evidence and no reminder vs. pre-evidence and reminder vs. post-evidence and no reminder vs. post-evidence and reminder vs. no-expert</td>
<td>Case strength: weak vs. strong (based on circumstantial evidence)</td>
<td>Skepticism (post-evidence and reminder)</td>
</tr>
<tr>
<td>Geiselman &amp; Mendez (2005)</td>
<td>None vs. expert testimony with closing arguments and judge instructions vs. closing arguments and judge instructions vs. attorneys closing arguments</td>
<td>WIC*: poor vs. good (i.e., lighting, distance, duration of view, emotional state of the witness, witness vision, cross-race identification, presence of disguise, suggestiveness of lineup instructions, showup or lineup, same day or two-day delayed identification)</td>
<td>Sensitivity (expert testimony)</td>
</tr>
<tr>
<td>Martire &amp; Kemp (2009)</td>
<td>None vs. congruent expert testimony vs. incongruent expert testimony vs. judicial instructions</td>
<td>Eyewitness accuracy: accurate vs. inaccurate</td>
<td>Null effect (all manipulations)</td>
</tr>
<tr>
<td>Jones et al. (2017)</td>
<td>None vs. expert testimony vs. <em>Henderson</em> instructions vs. enhanced instructions vs. expert testimony and <em>Henderson</em> instructions</td>
<td>Witnessing conditions: poor vs. good (i.e., exposure duration, weapon presence, delay) Identification conditions: poor vs. good (lineup type, lineup instructions, presence of feedback)</td>
<td>Skepticism (expert testimony) Null effect (both types of judicial instructions)</td>
</tr>
</tbody>
</table>

*Witnessing and identification conditions.
Overall, the results suggest that expert testimony has the potential to influence jurors’ judgments. However, the specific effects of expert testimony are inconsistent, with some studies showing skepticism and others showing limited sensitivity to witnessing and identification conditions. Expert testimony about the psychology of eyewitness identification does not address the case at hand and instead covers general issues of eyewitness memory and discusses the eyewitness literature without tailoring the content of the testimony to a particular case. This general nature of expert testimony might be responsible for the lack of a sensitivity effect.

**Comparative Analysis: Judicial Instructions vs. Expert Testimony**

Understanding the advantages and disadvantages of the two corrective safeguards can help establish which corrective safeguard should be the preferred method for educating the jury. In this section, we synthesize the legal and psychological perspectives by summarizing the advantages and disadvantages of both safeguards. The empirical evidence suggests that expert testimony is potentially more effective than judicial instructions because it had some influence on jurors’ judgments (i.e., skepticism), even though it did not consistently improve sensitivity. Table 4 summarizes the limited number of studies that included direct comparisons of judicial instructions and expert testimony (Cutler et al., 1990; Jones et al., 2017; Martire & Kemp, 2009).

**Advantages**

Judicial instructions require relatively little time to deliver and do not impose additional costs on the defendant (Sheehan, 2011; Simmonsen, 2011). In an adversarial system, and in particular in a jury trial, the judge’s role is limited in terms of the selection, presentation, and testing of evidence. Judges are impartial as to the outcome, while at the same time charged with the responsibility to ensure that the defendant receives a fair trial according to law. Importantly, judicial instructions carry with them the authority of the court and are less likely
to be perceived as aligned with the interests of either the prosecution or defense (Simmonsen, 2011). As compared to the use of expert testimony, this might decrease the probability that jurors rely too much on less relevant criteria (e.g., the expert’s credentials) and disregard eyewitness testimony on irrelevant grounds. This also means that jurors do not have to evaluate the credibility of another witness (which is the case with expert testimony) and can instead focus on the information provided in the instructions.

Another advantage is that it is easy to create uniform judicial instructions, whereas expert testimony will vary from one case (and one expert) to another. Moreover, uniform instructions could be crafted in advance to accurately reflect scientific findings and eliminate the possibility of misinforming the jury. And, at least in jurisdictions with uniform procedural guides (such as Judicial Bench Books, or Practice Notes issued by a Court), it should be easier to implement revised versions of judicial instructions to better reflect emerging scientific evidence because they are already a part of a trial and do not require any significant procedural changes (Simmonsen, 2011).

One of the main advantages of expert testimony is that, contrary to the vast majority of judicial instructions, the information provided by an expert is based on psychological research and is within the scope of the expert’s competence. As such, expert testimony is more likely to reflect the current state of knowledge. Moreover, experts usually provide explanations as to the magnitude and direction of eyewitness effects rather than simply listing them. Expert testimony is given in a question-and-answer format which might improve jurors’ ability to focus on the relevant information. Finally, experts are cross-examined, which gives jurors an opportunity to critically examine an expert’s claims (Dufraimont, 2008).
### Table 4

*Summary of Studies that Compared Judicial Instructions and Expert Testimony*

<table>
<thead>
<tr>
<th>Author</th>
<th>Safeguard</th>
<th>Evidence</th>
<th>Judicial instructions</th>
<th>Expert testimony</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cutler, Dexter, &amp; Penrod</td>
<td>None vs. expert testimony vs. <em>Telfaire</em> instructions.</td>
<td>WIC*: poor vs. good (i.e., presence of disguise, presence of a weapon, length of retention interval, lineup instructions)</td>
<td>Null effect</td>
<td>Skepticism</td>
</tr>
<tr>
<td>(1990)</td>
<td></td>
<td>Eyewitness confidence: 80% vs. 100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Martire &amp; Kemp</td>
<td>None vs. congruent expert testimony vs. incongruent expert testimony vs. judicial instructions</td>
<td>Eyewitness accuracy: accurate vs. inaccurate.</td>
<td>Null effect</td>
<td>Null effect</td>
</tr>
<tr>
<td>(2009)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jones et al.</td>
<td>None vs. expert testimony vs. <em>Henderson</em> instructions vs. enhanced instructions vs. expert testimony and <em>Henderson</em> instructions</td>
<td>Witnessing conditions: poor vs. good (i.e., exposure duration, weapon presence, delay) Identification conditions: poor vs. good (lineup type, lineup instructions, presence of feedback)</td>
<td>Null effect</td>
<td>Skepticism</td>
</tr>
<tr>
<td>(2017)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Witnessing and identification conditions.

**Disadvantages**

Both corrective safeguards have flaws. One of the most striking disadvantages of judicial instructions is that the evidence repeatedly demonstrates that they do not improve jurors’ ability to differentiate between reliable and unreliable eyewitness evidence. Beyond the empirical evidence, legal scholars have noted additional issues with judicial instructions. Given that the judge reads the judicial instructions, jurors might be influenced by his/her authority instead of critically evaluating the information (see Dufraimont, 2008). This might produce skepticism if jurors perceive the judge’s instructions as a critique of eyewitness evidence. Similarly, judicial instructions are not subject to cross-examination and, thus, the information is not open for explicit critical evaluation. Moreover, some judicial instructions...
contain inaccurate information and, as such, misinform the jury (Sheehan, 2011). For example, some judicial instructions appeal to jurors’ common sense when evaluating the reliability of eyewitness evidence, even though some eyewitness factors are unknown to lay people or are counterintuitive (Benton et al., 2006).

Judges usually have the discretion to modify judicial instructions by omitting what they deem to be unnecessary factors. This is problematic from an evidence-based perspective. Changing an existing set of previously-tested instructions invalidates any previous demonstrations of their effectiveness (if they existed). Finally, the judge usually reads the instructions at the end of the trial, after the presentation of all evidence and along with instructions on other matters, which might reduce their effectiveness (Ogloff & Rose, 2005; Ramirez et al., 1996; Sheehan, 2011).

Despite the empirical evidence, some legal scholars argue that judicial instructions could be effective if they are rewritten and presented before the evidence (e.g., Sheehan, 2011; Leverick, 2016), or if they are accompanied with visual aids (Simmonsen, 2011). Unfortunately, so far, empirical tests of improved instructions do not corroborate their optimism. The recent tests of the Henderson instructions have demonstrated that even when the instructions were rewritten to provide explanations, were presented before the evidence, and were provided to jurors as a handout, the instructions still did not sensitize jurors to the quality of eyewitness evidence (Dillon et al., 2017; Jones et al., 2017).

One of the most frequently mentioned disadvantages of expert testimony is that it is expensive and is, therefore, unavailable to many defendants (Sheehan, 2011; Simmonsen, 2011). Secondly, expert testimony can lead to a “battle of the experts” if the prosecutor and the defense attorney put forward experts who make opposing claims. This might confuse jurors, distract them from the actual evidence, and lead them to believe the expert who has better credentials rather than the one who makes the most convincing arguments (Sheehan,
Thirdly, experts often describe the results of the research without explaining how these results were obtained; thus, jurors must presume that the expert presented only valid research findings (Sheehan, 2011).

In addition, the information provided via instructions or expert testimony is general in nature. Neither the judicial instructions, nor the expert can or should comment on whether the specific witness testimony is unreliable, or whether the eyewitness evidence should be rejected. Therefore, even when the fact finder does comprehend the information provided in the educational aid, their knowledge of eyewitness factors at a general level does not necessarily translate into an effective evaluation of a particular eyewitness. The fact finder is always at liberty to reject the information provided as inapplicable in the current case. Finally, it is difficult to definitely establish the effectiveness of expert testimony because—in laboratory studies and in real cases—the testimony itself, and the expert, will vary. Even if one study establishes the effectiveness of the expert testimony, it does not necessarily guarantee that another expert’s testimony, even with similar content, but presented differently, will be as effective.

**Conclusions and Future Directions**

Legal systems around the world have recognized the frailties and weaknesses inherent in eyewitness testimony. Australia, Canada, and the USA have responded by developing (inconsistently applied) exclusionary rules, and, more recently, relying heavily on corrective safeguards, particularly judicial instructions. The legal safeguards discussed in this chapter have been influenced by empirical research. They have also, to some extent, incorporated aspects of this research into rules and procedures designed to guide investigative conduct, and the admission and reception of the evidence in court. As we have shown, there is little evidence that the existing safeguards are adequate to protect against risks that result from the admission of eyewitness evidence that might be unreliable either because of the
circumstances of the witnessing conditions or the suggestiveness of the identification procedures.

Overall, although the research shows that trial safeguards can influence jurors’ decision-making, the limited and inconsistent nature of this influence calls into question the extent to which legal systems should rely on these corrective safeguards as an alternative to excluding weak and potentially unreliable identification evidence. Corrective safeguards cannot definitively and consistently mitigate the damage created by admitting unreliable eyewitness evidence. Thus, we contend that courts should be more willing to exclude eyewitness evidence.

We must acknowledge that, despite incorporating some findings of empirical research into policies, legal scholars and practitioners still resist embracing psychological recommendations (e.g., Leverick, 2016). Some critics suggest that conclusions reached in most (if not all, in the eyewitness area) jury decision-making studies should be regarded with special caution and should not be extrapolated to real trials. One of the potential reasons for this resistance is a difference in the way legal scholars and psychologists evaluate the validity of jury decision-making studies. Psychological research relies heavily on simulations of trial experiences and research conducted with student participants, whereas legal scholars consider studies to be valid only if the study procedures are the same as those used during actual trials.

From a psychological perspective, evaluations of the validity of research findings should be informed by the available evidence; several of the criticisms articulated by legal scholars are not supported by empirical examinations of the variables concerned (e.g., no significant differences emerged between student and community samples in a meta-analysis of 53 jury decision-making studies, Bornstein et al., 2017). We acknowledge that the psychological studies do not capture the entirety of the courtroom experience and, thus, lack the degree of ecological validity desired by some legal scholars (Leverick, 2016). The
systematic and experimental psychological approach, however, has other strengths. Researchers can examine and isolate the influence of a number of individually-manipulated variables (e.g., suggestive identification procedures, expert testimony) on jurors’ decisions. Researchers examining a trial in its entirety would have less experimental control and, therefore, less understanding of the factors that ultimately influenced the verdict.

Nonetheless, psychologists must acknowledge that real cases will rarely be as clean-cut as those used in research studies. Likewise, legal scholars and practitioners must acknowledge that, although research fails to capture the complexity of real cases, they should not ignore the relevant empirical evidence. We highlight this debate, not to dismiss the concerns of either side, but rather to encourage both sides to acknowledge the strengths of different approaches and to advance the field through increased collaboration.

If we want courts to consider relevant empirical evidence when making their decisions, then researchers must provide them with such evidence. There are multiple gaps in our empirical knowledge. For example, we do not know: (a) if jurors process the same information provided by the judge and the expert in the same manner; (b) how deliberation affects the effectiveness of both corrective safeguards in eyewitness cases; and (c) why expert testimony often leads to skepticism, while judicial instructions fail to have any effect. Research in this area has remained relatively stagnant. Few efforts have been made to develop innovative ways of educating jurors about eyewitness issues (Wise, Fishman, & Safer, 2009 is a notable exception) and little is known about how video-recorded identification procedures will interact with legal safeguards. Furthermore, little attention has been drawn to new challenges posed by the proliferation of social media platforms (Skalon & Beaudry, in press) and contemporary developments in surveillance (e.g., Edmond & San Roque, 2013; McGorrery, 2015; San Roque, 2107). Even recent reports (e.g., NAS, 2014) have missed the significance of these technological developments in their analysis of the
challenges facing the field. Ultimately, a collaboration between psychologists and lawyers will provide evidence-based solutions to these unanswered questions and new challenges.
Acknowledgments

This work was partially supported by the Swinburne University Postgraduate Research Award.
Chapter 3. Method

In this chapter, I describe the methods and materials I used to conduct the four empirical studies. The main purpose of these studies was to investigate whether an educational aid (i.e., judicial instructions, expert testimony, or the Eyewitness Instructions and Checklist [EIC]) influences mock-jurors’ perceptions of eyewitness evidence. A literature review indicated that the common approach to conduct these studies is an experimental simulation of a trial. In my studies participants read or watched information about a fictional case, read or watched eyewitness evidence, read or watched a legal safeguard/educational aid (in experimental condition), and then completed a questionnaire.

Design

All four studies presented in this thesis used a between-subjects factorial design. The main independent variables were: the presence of one of the educational aids, identification accuracy, and the suggestiveness of identification procedures. In addition, in Paper 2, Study 2, I manipulated whether participants watched a video of an identification attempt and the eyewitness’s testimony or the eyewitness’s testimony alone. Table 5 outlines independent variables for each of the studies.

Independent Variables

Educational aids. In Paper 1, I manipulated the presence of the Victorian judicial instructions (cf. control). In Paper 2, Study 1, mock-jurors either watched the traditional expert testimony, the traditional expert testimony with visual aids, or completed a filler activity. In Paper 2, Study 2, mock-jurors watched the traditional expert testimony or completed a filler activity. Finally, Paper 3 tested the EIC (cf. control). More information about the educational aids is presented in the materials sections of this chapter and in corresponding chapters.
**Design Summary**

<table>
<thead>
<tr>
<th>Educational aid</th>
<th>Paper 1</th>
<th>Study 1</th>
<th>Study 2</th>
<th>Paper 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Traditional expert testimony vs expert testimony with visual aids vs control</td>
<td>Traditional expert testimony vs control</td>
<td>EIC vs control</td>
<td></td>
</tr>
<tr>
<td>Identification accuracy</td>
<td>Correct vs mistaken</td>
<td>Correct vs mistaken</td>
<td>Correct vs mistaken</td>
<td>Correct vs mistaken</td>
</tr>
<tr>
<td>Identification procedure</td>
<td>Non-suggestive vs. suggestive</td>
<td>Non-suggestive vs. suggestive</td>
<td>Non-suggestive vs. suggestive</td>
<td>Non-suggestive vs. suggestive</td>
</tr>
<tr>
<td>Evidence type</td>
<td>Video ID + testimony</td>
<td>Video ID + testimony</td>
<td>Video ID + testimony vs testimony only</td>
<td>Video ID + testimony</td>
</tr>
</tbody>
</table>

**Identification accuracy.** Half of the participants in each of the four studies watched an eyewitness who made a correct identification from a target-present lineup, whereas the other half watched an eyewitness who chose a filler from a target-absent lineup.

**Identification procedure.** I manipulated the suggestiveness of identification procedures. Non-suggestive identification procedures were identical in all four studies and included: double-blind lineup administration (neither the administrator nor the eyewitness knew the identity of the suspect), non-suggestive lineup instructions (the administrator warned the witness that the culprit may or may not be present), and a fair lineup construction (the suspect did not stand out from other lineup members). Suggestive identification conditions in Papers 1 and 3 were: single-blind lineup administration (the administrator knew the identity of the suspect), suggestive lineup instructions (i.e., “All you have to do is pick him out”), and a biased lineup (see Figure 1 for a sample fair lineup and a sample biased lineup).
Figure 1

A Sample Fair Lineup (Left) and a Sample Biased Lineup (Right)

Note. The yellow arrow point to the suspect; it was not present at the time of the identification, it was added for mock-jurors.

In Paper 2, in the suggestive identification condition the administrator of the lineup knew the identity of the suspect (i.e., single-blind administration) and provided confirmatory post-identification feedback by saying: “Good, you got the guy!” regardless of the actual accuracy of the eyewitness’s identification.

Evidence type. I manipulated this variable only in Paper 2, Study 2. Half of the participants watched a video-recorded identification procedure and the testimony of the eyewitness (same evidence type as in all other studies included in this thesis), whereas the other half watched the eyewitness’s testimony only.

Procedure

Participants signed up for the study and received a link to the online study (Qualtrics, Provo, UT). Following the link, participants read the consent information statement, provided their consent, and answered demographic questions. Paper 1 used a mini-trial paradigm in which participants read instructions on their decision-making role and the distinction between a suspect and a culprit, but there was no case summary, or opening and closing arguments.
from the attorneys. Papers 2 and 3 used a full-trial paradigm. Participants read (Papers 1 and 3) or watched (Paper 2) a case summary and attorneys’ opening arguments, watched eyewitness evidence, watched or read an educational aid (except Paper 3 where the EIC was presented before the evidence) and completed a questionnaire. Please refer to Figure 2 for the general outline of the procedure. A more detailed explanation can be found in the corresponding papers.

Figure 2

General Outline of the Procedure

| Participant signs up for the study |
| Information statement |
| Demographic questions |
| Paper 3 only |
| EIC (video) vs. control (video) |
| Case summary |
| Eyewitness evidence |
| Judicial instructions/expert testimony/expert testimony with visual aids/EIC (checklist)/control |
| Closing arguments |
| Questionnaire |
| Debriefing |

Materials

Case summary and attorneys’ arguments (Papers 2 and 3). Paper 2 employed a full-trial paradigm and presented mock-jurors with a video of the judge’s case summary and the
attorneys’ opening and closing statements. The case summary was provided by a person who played the judge; the opening statements by attorneys presented the facts of the case highlighting either strengths or weaknesses of the case. Please see Appendix 3 for the full trial transcript. Paper 3 used similar stimuli, however, instead of watching the videos, participants read the case summary and the attorneys’ statements (see Appendix 7).

**Eyewitness evidence stimuli.** The video-recorded eyewitness evidence that participants watched in the four studies was created for a larger project, part of which was reported in Beaudry et al. (2015). To create the videos, eyewitness-participants watched a mock-crime video and were asked to identify the culprit from a six-person simultaneous lineup and answer a series of questions about the crime, the culprit and the identification procedure. The identification procedure was manipulated to be suggestive or non-suggestive. The lineup was also manipulated to either contain the person they saw in the mock-crime video (i.e., a target-present lineup) or six fillers (i.e., a target-absent lineup). Correct eyewitnesses identified the target, whereas mistaken eyewitnesses identified a filler. Please see Figure 3 for a screen capture of a video-recorded identification procedure and the eyewitness’s testimony.

Figure 3

*Still Images of an Identification Procedure (Left) and the Eyewitness’s Testimony (Right)*

To increase stimulus sampling (Wells & Windschitl, 1999), the eyewitness evidence stimuli included 24 sets of eyewitness videos (two videos in each set: an identification procedure and the eyewitness’s testimony):
• Correct identification; double-blind administration (four eyewitnesses).
• Mistaken identification; double-blind administration (four eyewitnesses).
• Correct identification; single-blind lineup administration, suggestive lineup instructions, biased lineup (four eyewitnesses).
• Mistaken identification; single-blind lineup administration, suggestive lineup instructions, biased lineup (four eyewitnesses).
• Correct identification; single-blind lineup administration and post-identification feedback (four eyewitnesses).
• Mistaken identification; single-blind lineup administration and post-identification feedback (four eyewitnesses).

**Judicial instructions (Paper 1).** The Victorian judicial instructions list a number of eyewitness factors that jurors should consider (i.e., circumstances of observation, factors concerning the witness, and factors concerning the identification). See Appendix 2 for the full text of the instructions that was presented to participants. I used the text of Victorian judicial instructions as it is presented in the Victorian Criminal Charge Book, except for the sections concerning familiarity with the accused, quality of the material, and voice recognition, because these matters were not relevant to my study.

**Expert testimony stimuli (Paper 2).** I prepared the expert testimony script based on transcripts of real trials and in consultation with psychologists who have testified in the past. The expert testified in response to an examination by the defence attorney and by the prosecutor. The expert was recorded facing the camera; the attorneys asked the questions from off-camera. The expert first provided information about the three memory stages (i.e., encoding, retention, retrieval) and explained how eyewitness researchers conduct studies. Following this, the expert described factors that can influence the reliability of eyewitness evidence, and categorised them according to memory stage (i.e., factors at the encoding...
phase, factors at the retention phase, and factors at the retrieval stage). See Appendix 3 for the full text of the expert testimony (along with the judge’s case summary and attorneys’ arguments). In addition to this, Paper 2, Study 2 presented visual aids that appeared on the screen during relevant sections of the expert testimony (see Figure 4).

Figure 4
*A Slide Used in Expert Testimony with Visual Aids Condition, Accompanying Verbal Information about the Post-Identification Feedback*

Eyewitness Instructions and Checklist (Paper 3). The EIC consists of two parts: a video (narrated PowerPoint presentation) and a checklist. Similarly to the expert testimony, the video presentation first provided an overview of three memory stages and the methods used by eyewitness researchers; in addition, the video explained the principles of identification procedures and the differences between a suspect, a culprit, and a filler. The focus, however, was on eyewitness factors which were separated into three categories:

- Factors at the time of the crime: duration, distance, lighting, presence of a weapon, intoxication, disguises, stress, and cross-race identification.
• Factors between the criminal event and the identification procedure: delay and changed appearance.

• Factors at the time of the identification: lineup administration (single-blind or double-blind), lineup instructions, post-identification feedback, multiple viewings of the suspect, and the relationship between confidence and accuracy.

The second part—the EIC checklist—was presented after the evidence, but before the final questionnaire. The checklist asked participants to consider whether each of the factors was present in the case they watched and they could tick one of the following: “yes”, “no”, or “unknown” (see Figure 5).

Figure 5

Part of the Checklist from the EIC

<table>
<thead>
<tr>
<th>Factors at the time of the crime</th>
<th>Yes</th>
<th>No</th>
<th>Unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Duration. Did the witness briefly see the culprit?</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cross-race effects. Are the culprit and the witness of different races?</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Distance. Was the witness far away from the culprit?</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Lighting. Were the lighting conditions poor?</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Presence of weapon. Was a weapon present?</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Intoxication. Was the witness intoxicated?</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Disguises. Did the culprit wear a disguise during the crime?</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Stress. Was the event highly stressful?</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cross-race effects. Are the culprit and the witness of different races?</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
After completing the checklist, participants received the following instructions and could write their answer in a textbox (see Appendix 5 for the EIC transcript and Appendix 6 for the EIC checklist):

Please remember that it is up to you to decide whether and to what extent these factors affected the witness and to evaluate the identification evidence in combination with all of the evidence in this case. Keeping in mind the presence or absence of the impairing factors listed above, how would you evaluate the overall quality of eyewitness identification evidence in this case? Please explain in your own words.

**Control condition (Papers 1 and 2).** Mock-jurors in the control condition completed a filler activity. Participants were asked to memorise strings of random letters and numbers and recall them after a small delay (similar to the task used in Martire & Kemp, 2009).

**Control condition (Paper 3).** The educational aid (i.e., EIC) in this study consisted of two parts—for this reason activities in the control condition were also separated into two parts. In the beginning of the study, participants in the control condition watched a jury duty video and completed a brief questionnaire that tested their understanding of the information presented in the video. The second part was presented at the end of the study and acted as a control for the checklist. Participants were asked to describe factors that they thought were important in the case and the strengths and weaknesses of the case that they have identified.

**Questionnaire.** The main dependent variables in all four studies were verdict, confidence-in-verdict, the likelihood that the person identified by the witness was the culprit, perceptions of the witness, and perceptions of the lineup administrator. In addition to the main questionnaire, participants completed a questionnaire that tested their understanding of eyewitness factors. In Paper 1, only participants who read the judicial instructions completed this questionnaire. In all other papers, I compared the knowledge of eyewitness factors of
participants who were exposed to an educational aid and participants in the control condition. Please refer to Appendix 8 for the list of dependent variables.

**Manipulation checks.** Manipulation checks asked participants to describe what they saw in both eyewitness videos and to indicate whether they read/watched an educational aid or completed a filler activity. In addition, in Paper 3 participants completed simple attention checks (e.g., $2 + 2 = ?$).

**Data Analytic Approach**

I used a similar approach in all four studies. I analysed the binary verdict measure with a binary logistic regression, with the main effects and all interactions of the three independent variables. I used ANOVAs to examine the independent and interactive effects of educational aid, identification accuracy, identification procedure, and evidence type (Paper 2, Study 2 only) on the confidence-in-verdict, the likelihood that the person identified by the witness was the culprit, perceptions of lineup fairness, perceptions of the lineup administrator, and perceptions of the eyewitness. I used chi-square analyses to examine the effects of educational aids on participants’ knowledge of eyewitness factors. Finally, I also examined continuous dependent variables using a Bayesian analysis to investigate whether the findings of null-hypothesis significance testing would be supported (JASP Team, 2018).
Chapter 4. Paper 1

Original Research

The Effectiveness of Judicial Instructions on Eyewitness Evidence in Sensitizing Jurors to Suggestive Identification Procedures Captured on Video

Alena Skalon and Jennifer L. Beaudry

Swinburne University of Technology

Author note: Alena Skalon, Department of Psychological Sciences; School of Health Sciences; Faculty of Health, Arts and Design; Swinburne University of Technology, Hawthorn, Victoria, Australia; email: askalon@swin.edu.au.

Jennifer L. Beaudry, Department of Psychological Sciences; School of Health Sciences; Faculty of Health, Arts and Design; Swinburne University of Technology, Hawthorn, Victoria, Australia; email: jbeaudry@swin.edu.au.

Please note that I have made small alterations to this Paper for thesis presentation. These changes were limited to placement of tables and figures, numbering of tables and figures, removing all reference lists (which are compiled at the end of the thesis), and altering some in-text citations to align with this compiled list. As no other changes were made, the referencing style and language requested by the *Journal of Experimental Criminology* were retained.
Abstract

Objectives. Jurors may believe eyewitnesses even when their identifications are obtained from suggestive administration conditions. One of the legal safeguards designed to educate jurors about eyewitness evidence is judicial instructions. However, their effectiveness in sensitizing jurors to genuine eyewitness evidence captured on video is unknown.

Methods. Mock-jurors ($N = 232$) watched the video-recorded identification and testimony of one of 16 genuine eyewitnesses. We varied the suggestiveness of the identification procedure, whether they saw an accurate or inaccurate identification, and whether they received judicial instructions about eyewitness evidence.

Results. Mock-jurors were sensitive to eyewitness accuracy when an identification procedure was non-suggestive, rendering more guilty verdicts when the eyewitness was accurate than inaccurate; this sensitivity was impaired when an identification was made under suggestive circumstances. The judicial instructions manipulation did not significantly affect verdicts. Judicial instructions did, however, lead to confusion: participants who read the judicial instructions (cf. control) were more willing to believe the witness when the identification was obtained under suggestive circumstances.

Conclusions. Suggestive identification procedures impaired mock-jurors’ sensitivity to eyewitness accuracy and the judicial instructions failed to improve their sensitivity. This was the first study to our knowledge to demonstrate the ineffectiveness of judicial instructions when mock-jurors viewed both eyewitness testimony and video-recorded identification procedures. We also included a Bayesian analysis to investigate whether the findings of null-hypothesis significance testing would be supported. Traditional legal safeguards may not be an effective remedy for the damaging effects of suggestive identification procedures.

Keywords: eyewitness evidence, judicial instructions, juror decision making, jury instructions, video-recorded identification procedure.
The Effectiveness of Judicial Instructions on Eyewitness Evidence in Sensitizing Jurors to Suggestive Identification Procedures Captured on Video

Eyewitnesses often play a crucial role in the trial process because they can shed light on one of the main issues in criminal cases—the identity of the perpetrator. Despite the utility of eyewitness testimony, real cases and psychological research reveal that eyewitnesses can sometimes mistakenly identify innocent people as perpetrators (e.g., The Innocence Project 2018). Importantly, the misidentification itself does not lead to a conviction. It is up to the fact-finders (e.g., judge or jury) to evaluate the entirety of the evidence, including the eyewitness testimony, and decide the guilt of the defendant. Unfortunately, people often believe eyewitnesses regardless of the actual quality of the witnessing and identification conditions or the accuracy of the identification (for a review, see Boyce et al. 2007). The present study examines whether judicial instructions can sensitize mock-jurors to suggestive identification procedures and to eyewitness accuracy when mock-jurors viewed a genuine-eyewitness’s video-recorded identification procedure and testimony.

**Video-recorded Identification Procedures**

Video-recording the identification procedure itself may be one of the ways to overcome the overbelief of eyewitness testimony (Kassin 1998; Reardon and Fisher 2011; Wilford and Wells 2013). It may provide useful information about the confidence of an eyewitness at the time of their identification, rather than during their testimony (Brewer and Wells 2006; Wixted and Wells 2017) and preserve informative cues that may otherwise be obscured by the eyewitness’s preparation for a trial (e.g., Reardon & Fisher 2011; Kaminski & Sporer 2017).

Indeed, when Reardon and Fisher (2011) exposed mock-jurors to the video-recorded identification procedure along with the eyewitness testimony (cf. testimony alone), mock-jurors were better able to differentiate accurate from inaccurate eyewitnesses. In other words,
they showed sensitivity to eyewitness accuracy when they watched the eyewitness make their identification decision from a simultaneous lineup. Beaudry et al. (2015) also found evidence of sensitivity to eyewitness accuracy, but only when evaluators saw eyewitnesses make their identification from simultaneous, but not sequential, lineups. Moreover, sensitivity to eyewitness accuracy was eliminated when mock-jurors viewed eyewitnesses who received post-identification feedback (Beaudry et al. 2015). These results suggest that jurors who view video-taped identification procedures may be able to discriminate between accurate and inaccurate eyewitnesses, but perhaps only when the evidence was obtained from non-suggestive, simultaneous identification procedures. Douglass and Jones (2011) tested whether presenting video-recorded identification procedure (cf. reading confidence statement) sensitized participants to confidence inflation. Their findings indicated that participants were more likely to lower ratings of the defendant’s guilt when they watched video-recorded identification procedure (cf. hearing confidence statement read by the witness) and there was a discrepancy between the witness’s confidence at the time of the identification and at the time of the trial. In another study, mock-jurors who watched a scripted video-taped identification procedure and the eyewitness testimony (cf. testimony only) demonstrated greater sensitivity to the suggestiveness of single-blind administration (Modjadidi and Kovera 2018). Critically, neither manipulated eyewitness accuracy, so it is unclear as to whether their mock-jurors were sensitive to accuracy. Smalarz and Wells (2014) manipulated the presence of post-identification feedback and eyewitness accuracy and reached similar conclusion: jurors’ ability to discriminate between accurate and inaccurate eyewitnesses was impaired when they watched testimony of eyewitnesses who received post-identification feedback.

**Knowledge of Eyewitness Factors**

One of the possible explanations for the decrease in sensitivity to accuracy when presented with evidence from suggestive identification procedures is jurors’ lack of
knowledge about factors that influence the reliability of eyewitness evidence. When lay people are asked to list the factors that can affect eyewitness accuracy, they tend to focus on factors that were present at the time of the crime (e.g., lighting), the characteristics of the culprit (e.g., race and physical appearance), or the characteristics of an eyewitness (e.g., vision), but rarely spontaneously mention the procedures used by the police (Shaw et al. 1999). One comparison of experts in eyewitness psychology and jurors revealed that the two groups disagreed on 87% of the 30 items, including eight items assessing system variable knowledge (e.g., lineup instructions, post-event information; Benton et al. 2006). On the other hand, a meta-analysis of 23 surveys demonstrated that lay people disagreed with experts on only 29% of questions that assessed knowledge regarding system variables (Desmarais and Read 2011). Despite such a high rate of agreement, Desmarais and Read (2011) warn that knowledge of eyewitness factors does not necessarily translate into an ability to make accurate judgments in real court settings. Indeed, in a simulated court case, jurors appropriately rated suggestive procedures as less fair compared to non-suggestive procedures, but that awareness did not lead to significantly fewer verdicts in the suggestive condition (Devenport et al. 2002).

**Judicial Instructions**

One of the legal safeguards used to educate jurors about the relevant factors that they should consider when evaluating eyewitness evidence is instructions that the judge provides to the jury (i.e., judicial instructions). Through these instructions, judges present—usually at the end of the trial—information about the psychology of eyewitness identifications and discuss a number of factors that should help jurors assess this evidence. Judicial instructions, like any other educational or corrective safeguard (e.g., expert testimony), can produce one of three effects: confusion, skepticism, or sensitivity (Cutler et al. 1989). Confusion means that jurors did not appropriately use the information provided by the instructions and, thus,
incorrectly evaluating the evidence. For example, after hearing judicial instructions, they may give more weight to unreliable than reliable evidence. Skepticism occurs when jurors discount and disbelieve the eyewitness evidence regardless of the quality of the witnessing and identification conditions or the eyewitness’s accuracy. Finally, there are two types of sensitivity: sensitivity to witnessing and identification conditions and sensitivity to eyewitness accuracy (Martire and Kemp 2011). When jurors are sensitive to witnessing and identification conditions, they take into account the circumstances under which a witness viewed a crime and made an identification, believing an eyewitness only when the circumstances were favorable (i.e., good viewing conditions and non-suggestive identification procedures). When jurors are sensitive to eyewitness accuracy, they believe accurate eyewitnesses and disbelieve inaccurate ones (which, ideally, will then translate into guilty verdicts for culpable defendants and acquittals for innocent suspects).

Martire and Kemp (2009) underscore the difference between studies that measure sensitivity to witnessing and identification conditions and those that investigate sensitivity to eyewitness accuracy. In the first type, researchers manipulate the quality of witnessing and/or identification conditions in a transcript of a simulated case or they show a video of an eyewitness who is played by an actor (e.g., Devenport et al. 2002; Modjadidi and Kovera 2018). These studies can provide useful information about whether legal safeguards make jurors more knowledgeable about eyewitness factors; however, this knowledge does not necessarily translate into real-life settings. In real cases, jurors do not simply evaluate witnessing and identification conditions, but rather evaluate the eyewitness testimony as a whole. In contrast, studies that measure sensitivity to eyewitness accuracy are more ecologically valid. These studies use the genuine eyewitness paradigm, in which naïve participants first witness a crime, make identification decisions (where they may accurately identify the culprit from the mock-crime video or mistakenly identify an innocent lineup
member), and then testify about their experiences like they would in a real case (Wells et al. 1979). Ideally, at trial, jurors should be sensitive to eyewitness accuracy, believing accurate eyewitnesses and disbelieving inaccurate eyewitnesses. Few studies have manipulated both identification conditions and eyewitness accuracy (cf. Beaudry et al. 2015) and none, to our knowledge, have investigated the effects of legal safeguards in improving evaluations of this eyewitness evidence.

Having discussed the various possible effects of judicial instructions and how different paradigms are needed to investigate the two types of sensitivity, we now address whether judicial instructions are effective in sensitizing jurors to the quality of eyewitness evidence. Most studies examining this question manipulated witnessing and identification conditions and used US-based judicial instructions (e.g., US v. Telfaire 1972; New Jersey v. Henderson 2011). Telfaire instructions ask jurors to consider a number of factors that can affect the reliability of eyewitness evidence, for example, “If the identification by the witness may have been influenced by the circumstances under which the defendant was presented to him for identification, you should scrutinize the identification with great care” (Greene 1988: 3). Researchers criticized the Telfaire instructions for being poorly written and misleading (e.g., Cutler and Penrod 1995); thus, it is not surprising that these instructions did not increase sensitivity and instead lead to skepticism (Greene 1988), confusion (Ramirez et al. 1996: Experiment 1), or had no effect on jurors’ judgments (Cutler et al. 1990).

Unfortunately, attempts to improve judicial instructions by incorporating psychological research—the Henderson instructions—have not fared better than their predecessor. Henderson instructions either lead to skepticism (Papailiou et al. 2015; Dillon et al. 2017) or had no effect on jurors’ verdict (Jones et al. 2017). In these studies, jurors were presented with simulated eyewitness testimony. Thus, little is known about whether judicial instructions can influence jurors’ evaluation of genuine eyewitnesses when they view video-recorded
identification evidence and their accompanying testimony. The only study to date that tested the effectiveness of judicial instructions (Judicial Commission of NSW 2006) in sensitizing jurors to eyewitness accuracy produced a null effect (Martire and Kemp 2009).

**Present Study**

The present study evaluates the effectiveness of current Victorian judicial instructions (Judicial College of Victoria, Australia 2006) in improving mock-jurors’ sensitivity to identification conditions and eyewitness accuracy when presented with video-recorded identification evidence and testimony. The Victorian instructions inform the jury about the dangers of identification evidence and discuss the circumstances of observation, factors concerning the witness, and factors concerning the identification procedure. A possible limitation of these instructions is that the majority of these factors are presented as a question—the instructions ask the jury to consider a number of factors without explaining how they can affect the accuracy of the identification. Thus, given that laypeople may have a limited understanding of how various factors affect eyewitnesses, it is possible that the jury may misunderstand or miscalculate the effect of one or more factors on the eyewitness’s identification.

Participants watched eyewitness evidence (video-recorded identification and testimony) of either accurate or inaccurate eyewitnesses obtained under suggestive or non-suggestive identification conditions. Half then either read the judicial instructions and half completed a control activity. We hypothesized an Identification Accuracy × Identification Procedure interaction: participants will be sensitive to eyewitness accuracy when obtained from a non-suggestive (but not suggestive) identification procedure (Hypothesis 1). We also hypothesized that the judicial instructions may sensitize participants to the identification procedure conditions, producing a significant two-way Identification Procedure × Judicial Instructions interaction. Specifically, participants who read judicial instructions would
believe eyewitness identifications from non-suggestive conditions more than eyewitness identifications from suggestive conditions, whereas the control group would believe eyewitnesses regardless of the quality of identification procedure (Hypothesis 2). There is no evidence to suggest that judicial instructions may sensitize participants to eyewitness accuracy, but a significant two-way Identification Accuracy × Judicial Instructions interaction would lend support to this hypothesis (Hypothesis 3). That is, participants in the judicial instructions (cf. control) condition will believe accurate eyewitnesses more often than inaccurate ones, regardless of the quality of identification conditions.

**Method**

**Disclosures**

The data (both raw and clean) and the complete survey are available on the Open Science Framework (OSF; https://tinyurl.com/ydghqyvj). We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study. This project has been approved by Swinburne’s Human Research Ethics Committee (SUHREC) in line with the Australian National Statement on Ethical Conduct in Human Research.

**Design**

This study used a 2 (Identification [ID] Procedure: suggestive vs. non-suggestive) × 2 (ID Accuracy: accurate vs. inaccurate) × 2 (Judicial Instructions: Victorian judicial instructions vs. control) between-subjects factorial design. In the non-suggestive identification procedure condition, the lineup administrator did not know which lineup member was the target in the mock-crime video (i.e., was blind to the suspect), the administrator provided unbiased instructions warning that the culprit may or may not be present in the lineup, and the lineup was constructed to be fair with no lineup member standing out from the others (our mock-juror participants saw the lineup when they completed the study). In the suggestive identification procedure condition, the administrator
knew the identity of the suspect, provided instructions saying, “All you have to do is pick him out,” and the lineup was biased such that the suspect stood out from the lineup members. Accurate eyewitnesses correctly identified the target from the target-present lineup, whereas inaccurate eyewitnesses identified the innocent suspect from the target-absent lineup.

Participants

Participants were 232 (50 males, 182 females) first-year psychology and criminology students at a large Australian university who participated in exchange for extra credit. Participants’ ages ranged from 16 to 73 years ($M = 31.81, SD = 10.83$). Most participants self-identified as Caucasian (81%), followed by Asian (9%), “other” (9%), and Indigenous (0.9%). This study is a part of a PhD thesis, thus, the sample size was limited by the available data collection timeframe.

Materials

Witness identification and eyewitness testimony stimuli. This study used a selection of the stimuli created by Beaudry et al. (2015), selected based on their inclusion in the conditions detailed below. In each of the four conditions, we included identification and testimony stimuli from four eyewitnesses to increase stimuli sampling (Wells and Windschitl 1999). To create the stimuli, Beaudry et al. (2015) had participants individually view one of four targets in a mock-crime video (theft from an unattended purse). After a short delay, the participants were informed that they were eyewitnesses to a crime (hereafter referred to as eyewitnesses) and asked to participate in an identification procedure and to provide testimony. The identification procedure was video-recorded so that the administrator and the eyewitness were seen in profile sitting at the table. The stimuli in the present study showed eyewitness being presented with a six-person photographic simultaneous lineup of white males, that either contained the target (target-present) or an innocent suspect (target-absent). In all conditions, after the eyewitness made an identification decision, the administrator asked
the eyewitness to report their confidence in their identification in their own words. After a short delay, eyewitnesses were asked a series of testimony-relevant questions concerning the crime and the identification procedure. The testimony was video-recorded with the eyewitness facing the camera, and the researcher asked the questions from off-camera.

**Judicial instructions.** The Victorian judicial instructions (see Appendix 2) warn the jury that the identification evidence is potentially unreliable and ask jurors to consider factors regarding the circumstances of observation, factors concerning the witness, and factors concerning the identification. Additionally, the instructions advised that an identification from photographs rather than identification parades (i.e., live lineups) may be unreliable. The judicial instructions took approximately four minutes to read.

**Control.** Participants in the control condition completed a filler activity, which included memory tasks that took approximately two minutes to complete. Participants were asked to memorize strings of random letters and numbers and recall them after a small delay (similar to the task used in Martire and Kemp 2009).

**Questionnaire.** Please refer to Appendix 8 for the complete questionnaire. Our main dependent variables were verdict (binary: guilty vs. not guilty); confidence-in-verdict, ranging from -100% (extremely confident in the not guilty verdict) to 100% (extremely confident in the guilty verdict); and the likelihood that the person identified by the witness was the culprit, ranging from 0% (very unlikely) to 100% (very likely). Another key set of dependent variables included fairness of lineup presentation, ranging from 0 (completely unfair) to 10 (completely fair) and two combined measures. We averaged three questions to create the perceptions of the witness scale: the witness’s view of the culprit, attention the witness paid to the culprit’s face, and the clarity of the witness’s memory at the time of the lineup, all $rs > .65, ps = < .001$. This scale ranged from 0 (not good) to 10 (very good). We also averaged four questions to create the perceptions of the administrator’s influence scale: the officer’s
pressure to choose someone, the officer’s influence on the witness’s confidence, the officer’s knowledge of the suspect’s identity, and the officer’s investment in the outcome of the lineup, all \( r > .39, p < .001 \). This combined measure ranged from 0 (no influence) to 10 (definite influence). Participants also completed a belief measure (“Do you believe the person identified by the witness is actually the culprit?” binary: yes or no) and confidence-in-belief. The pattern for these measures was identical to verdict and confidence-in-verdict, for this reason, we do not report these results in this manuscript. Data from all questions are available https://tinyurl.com/y79j8afn

**Comprehension of judicial instructions.** In the Victorian judicial instructions (but not the control) condition, we measured participants’ comprehension of the instructions (e.g., Greene 1988; Jones et al. 2017). Judicial instructions cannot be effective if jurors do not understand them (e.g., Semmler and Brewer 2002); thus, it is important to assess whether participants understood the information presented to them. The comprehension questions included six multiple choice and four true/false questions (see Table 6) that covered many of the topics discussed in the judicial instructions (e.g., “The degree of similarity between the suspect and other people in the photoboard can influence the accuracy of an identification”).

**Manipulation checks.** Manipulation checks asked participants to describe what they saw in both eyewitness videos and to indicate whether they read the instructions or completed a filler activity. The full questionnaire is available at https://tinyurl.com/ybf7p5vt

**Procedure**

Participants signed up for the study and received a link to the online study (Qualtrics, Provo, UT). Following the link, participants read the consent information statement and provided their consent, answered demographic questions, and read instructions on their decision-making role and the distinction between a suspect and a culprit. Next, they saw a video-recorded identification procedure that was suggestive or non-suggestive and testimony
of an eyewitness who made an accurate or inaccurate identification. Following the videos, participants saw either a fair photo lineup or a biased photo lineup. After the eyewitness evidence, half of the participants read the Victorian judicial instructions, whereas the other half completed a filler activity. Participants then completed the questionnaire. At the end of the study, participants were thanked for their participation and debriefed.

Results

Data Preparation

Manipulation checks. The initial sample consisted of 336 participants. We excluded participants who did not finish the study ($n = 20$), who took either less than 10 or more than 55 minutes ($n = 37$) to complete the study ($M = 25.41$, $SD = 8.48$), who indicated that they did not see one of the videos ($n = 24$), and those who failed the instruction condition manipulation check ($n = 23$). The final sample consisted of 232 participants. Due to a programming error, 59 of these 232 participants did not complete the verdict and the confidence-in-verdict measures, but we report their data for all other measures.

Confidence-in-verdict measure. A confidence-in-verdict measure was created by multiplying the confidence rating by -1 in cases of a not-guilty verdict (Tenney et al. 2007). Thus, -100% indicates that a participant was very confident in a not-guilty verdict, whereas 100% indicates high confidence in a guilty verdict.

Data Analytic Approach

Each participant saw one witness, but we had four witnesses per condition. Rather than analyze each of the four witnesses as separate stimuli, we collapsed data across witnesses in the conditions (e.g., Beaudry et al. 2015). For example, we combined data from all four videos with non-suggestive identification procedures and accurate eyewitnesses. For our binary measure (verdict), we used a binary logistic regression, with the main effects and all interactions of three independent variables entered hierarchically (described in the
corresponding section). We report percentages for the binary measure. For our continuous measures (confidence-in-verdict; likelihood that the person identified by the witness was the culprit; perceptions of the lineup presentation; perceptions of the lineup administrator; and perceptions of the witness), we used ANOVAs to examine the independent and interactive effects of ID accuracy, ID procedure, and judicial instructions. For all continuous measures we report means and standard deviations. We report 95% confidence intervals in square brackets. In addition, we also conducted a Bayesian analysis to estimate the Bayes factor (JASP Team 2018) for the judicial instructions. Bayesian analysis compares the evidence in favour of both the null and the alternative hypotheses and quantifies the evidence (e.g., Wagenmakers et al. 2018; Jarosz and Wiley 2014).

**Verdict**

Variables were entered hierarchically; ID accuracy was entered in the first step, ID procedure was entered in the second step, judicial instructions manipulation was entered at the third step, ID Accuracy × ID Procedure interaction was entered in the fourth step, ID Accuracy × Judicial Instructions interaction was entered in the fifth step, ID Procedure × Judicial Instructions interaction was entered in the sixth step, and ID Accuracy × ID Procedure × Judicial Instructions interaction was added in the seventh step. The last three steps did not significantly improve the model, $p$s $>.42$, and were removed. The model that included three main effects and ID Accuracy × ID Procedure interaction was statistically significant, $\chi^2(4) = 25.99, p < .001$, explaining 19% (Nagelkerke $R^2$) of the variance in verdict, and correctly classifying 65.90% of cases. The judicial instructions main effect did not improve the model fit; thus, the final model included ID accuracy, ID procedure, and the ID Accuracy × ID Procedure interaction.

The final model was statistically significant, $\chi^2(3) = 24.93, p < .001$, explaining 18.10% (Nagelkerke $R^2$) of the variance in verdict, and correctly classifying 65.30% of cases. There
was a significant ID Accuracy × ID Procedure interaction, $B = 1.55$, $SE = .72$, Wald $(1, n = 173) = 4.71, p = .03$ (see Figure 6). When the ID procedure was non-suggestive, mock-jurors demonstrated sensitivity to eyewitness accuracy, with more mock-jurors rendering guilty verdicts after watching an accurate ID than an inaccurate ID (57% and 12%, respectively), $\chi^2(1) = 19.08, p < .001$, OR = 9.74 [3.22, 29.49], providing support for Hypothesis 1. However, when they viewed a suggestive ID procedure, mock-jurors’ verdicts did not significantly differ for accurate and inaccurate IDs (53% and 36%, respectively), $\chi^2(1) = 2.73, p = .10$, OR = 2.06 [0.87, 4.87]. This reduction in discrimination was driven by an increased belief of inaccurate IDs obtained under suggestive conditions rather than non-suggestive conditions, $\chi^2(1) = 6.56, p = .010$, OR = 0.24 [0.08, 0.75].

**Figure 6.** Two-way interaction between identification accuracy and identification procedure on mock-jurors’ verdicts. Error bars reflect 95% confidence intervals.

*Note.* *p < .05. **p < .001.
Confidence-in-Verdict Measure

Confidence-in-verdict ranged from -100 to 92. The ANOVA revealed a main effect of ID Accuracy, $F(1, 165) = 18.38, p < .001, d = 0.67 [-8.27, 9.61]$. Mock-jurors were more confident in guilty verdicts when an eyewitness was accurate ($M = 8.15, SD = 64.42$) and more confident in not-guilty verdicts when an eyewitness was inaccurate ($M = -31.98, SD = 55.67$).

There was a significant ID Accuracy × ID Procedure interaction on mock-jurors’ confidence-in-verdicts, $F(1, 165) = 5.10, p = .025, \eta^2_p = .03$. The interaction demonstrated that mock-jurors’ reported confidence-in-verdict was sensitive to eyewitnesses’ accuracy when the ID procedure was non-suggestive (accurate ID: $M = 9.75, SD = 64.98$; inaccurate ID: $M = -51.19, SD = 43.22$), $t(84) = 5.10, p < .001, d = 1.11 [-10.47, 12.69]$, but sensitivity was impaired when it was suggestive (accurate ID: $M = 6.58, SD = 64.55$; inaccurate ID: $M = -12.76, SD = 60.41$), $t(85) = 1.44, p = .15, d = 0.31 [-12.69, 13.31]$ (see Figure 7). The judicial instructions did not have significant independent or interactive effects on mock-jurors’ verdicts, all $F$s < 1.31, $ps > .25$. Please note that these large confidence intervals are caused by large SDs and the fact that the scale ranges from -100 to 100.

Likelihood that the Person Identified by the Witness was the Culprit

Mock-jurors estimated higher likelihood that the eyewitness correctly identified the culprit when the eyewitness made an accurate ID ($M = 55.78\%, SD = 21.46$) compared to an inaccurate ID ($M = 43.12\%, SD = 22.03$), $F(1, 224) = 21.93, p < .001, d = 0.59 [-2.20, 3.37]$. Mock-jurors also estimated a greater likelihood that the eyewitness correctly identified the culprit when the ID procedure was suggestive ($M = 53.34\%, SD = 23.21$) compared to non-suggestive ($M = 45.78\%, SD = 21.41$), $F(1, 224) = 8.88, p = .003 d = 0.34 [-2.52, 3.20]$. 
Figure 7. Two-way interaction between identification accuracy and identification procedure on mock-jurors’ confidence-in-verdict. Values above zero indicate confidence in a guilty verdict, whereas values below zero indicate confidence in a not-guilty verdict. Error bars reflect 95% confidence intervals.

These main effects should be interpreted within their significant interaction, $F(1, 224) = 12.46, p = .001, \eta_p^2 = .05$. Mock-jurors demonstrated sensitivity to eyewitness accuracy using the likelihood scale when the ID procedure was non-suggestive, but not suggestive, supporting Hypothesis 1. As shown with confidence-in-verdict, when mock-jurors viewed identifications obtained from non-suggestive procedures, mock-jurors estimated a greater likelihood that the eyewitness correctly identified the culprit when they viewed an accurate ID ($M = 56.68\%, SD = 18.24$) rather than an inaccurate ID ($M = 34.09\%, SD = 18.25$), $t(114) = 6.67, p < .001, d = 1.25 [-2.04, 4.54]$. When they viewed a suggestive procedure, mock-jurors’ estimates that the witness likely identified the culprit did not significantly vary according to ID Accuracy (accurate ID: $M = 54.84\%, SD = 24.48$; 58 inaccurate ID: $M = 51.84\%, SD = 21.99$), $t(114) = 0.69, p = .49, d = 0.13 [-4.07, 4.33]$.

A significant ID Procedure × Judicial Instructions interaction indicated that the judicial instructions lead to confusion, $F(1, 224) = 4.47, p = .036, \eta_p^2 = .02$ (see Figure 8). When the
ID procedure was suggestive, mock-jurors who read the judicial instructions (cf. control) estimated a greater likelihood that the eyewitness made an accurate ID (judicial instructions: \(M = 57.62\%, SD = 21.92\); control: \(M = 49.35\%, SD = 23.85\)) \(t(114) = 1.94, p = .055, d = 0.36 [-3.78, 4.50]\). This result is contrary to Hypothesis 2. There were no differences between the two instruction conditions when the procedure was non-suggestive, \(t(110) = -0.55, p = .59\).

Figure 8. Two-way interaction between identification procedure and judicial instructions on mock-jurors’ estimate of the likelihood that the witness made an accurate identification. Error bars reflect 95% confidence intervals.

**Bayesian ANOVA.** A three-way Bayesian ANOVA (JASP Team, 2018) demonstrated that the data were 343,358 times more likely to occur when the ID Accuracy \(\times\) ID Procedure interaction was included, \(BF_{10} = 343,358\) (cf. null model). Interestingly, the addition of Judicial Instructions to the model as an interaction term decreased the Bayes factor from 343,358 to 20,891, \(BF_{10} = 684.31\); thus, the data provided stronger support for a model that included only the ID Accuracy \(\times\) ID Procedure interaction.
Contrary to inferential statistics, a Bayesian ANOVA indicated that the data were 1.60 times more likely to occur under the null hypothesis when ID Procedure × Judicial Instructions interaction was included (cf. null model), BF\(_{01}\) = 1.60. Overall a Bayesian ANOVA suggests that the model that includes the ID Accuracy × ID Procedure interaction outperforms all other models. Moreover, when only the Judicial Instructions main effect was included in the model, the data were 3.96 times more likely to occur under the null hypothesis, BF\(_{01}\) = 3.96. The Bayesian analysis demonstrates that the judicial instructions did not affect mock-jurors’ perceptions of eyewitness accuracy.

**Perceptions of the Lineup Presentation**

The main effect of the ID procedure revealed that mock-jurors were aware of the biased nature of the suggestive procedure, \(F(1, 224) = 31.57, p < .001, d = 0.76 [0.45, 1.06]\). Mock-jurors rated non-suggestive ID procedures as more fair (\(M = 7.89, SD = 1.99\)) than suggestive procedures (\(M = 6.09, SD = 2.74\)). All other main effects and interactions were not significant, all \(Fs < 1.46, ps > .23\).

**Bayesian ANOVA.** A Bayesian ANOVA supported the inferential results; a model that included only ID procedure outperformed any other model (BF\(_{10}\) = 342,680). A Bayesian ANOVA demonstrated that a model that included a main effect of judicial instructions indicated that the data were 3.26 times in favour of the null hypothesis, BF\(_{01}\) = 3.26.

**Perceptions of the Lineup Administrator**

There was a main effect of ID procedure, \(F(1, 222) = 24.30, p < .001, d = 0.64 [0.36, 0.92]\) and a main effect of ID accuracy, \(F(1, 222) = 11.68, p = .001, d = 0.44 [0.16, 0.73]\). Mock-jurors perceived the lineup administrator as exerting more influence on the eyewitness when the ID procedure was suggestive (\(M = 4.03, SD = 2.40\)) than when it was non-suggestive (\(M = 2.64, SD = 1.95\)). Mock-jurors also reported greater administrator influence when they saw an inaccurate ID (\(M = 3.84, SD = 2.16\)) compared to an accurate ID (\(M = \ldots\)
2.85, $SD = 2.32$). Higher numbers indicate greater administrator influence. The judicial instructions did not significantly affect mock-jurors’ perceptions of the lineup administrator, all $Fs < 3.36, ps > .07$.

**Bayesian ANOVA.** The Bayes factor supported this result—under a model with the two main effects the data were 212,638 times more likely to occur, $BF_{10} = 212,638$ (cf. null model). When only the main effect of judicial instructions was included in the model—the data were 1.66 times in favour of the null hypothesis, $BF_{01} = 1.66$.

**Perceptions of the Witness**

A main effect of ID procedure indicated that mock-jurors reported more positive perceptions of the witness when the ID procedure was suggestive ($M = 5.91$, $SD = 1.93$) than non-suggestive ($M = 5.14$, $SD = 2.46$), $F(1, 224) = 7.60, p = .006, d = 0.35 [0.07, 0.63]$. A main effect of ID accuracy revealed that mock-jurors were sensitive to eyewitness accuracy, reporting a more positive perception of the witness when the ID was accurate ($M = 6.08$, $SD = 2.13$) than inaccurate ($M = 4.96$, $SD = 2.21$), $F(1, 224) = 15.56, p < .001, d = 0.52 [0.24, 0.80]$. Higher numbers indicate more positive perceptions of the witness. No significant interactions emerged and the judicial instructions did not significantly influence the perceptions of the witness, all $Fs < 1.87, ps > .10$.

**Bayesian ANOVA.** However, a Bayes factor provided slightly stronger support for an ID procedure × ID accuracy interaction, compared to two main effects ($BF_{10} = 1,447$ and $BF_{10} = 977$, respectively). A Bayesian ANOVA demonstrated that a model that included a main effect of judicial instructions indicated that the data were 2.50 times in favour of the null hypothesis ($BF_{01} = 2.50$).

**Comprehension of Instructions**

The questions and correct answers are reported in Table 6. For questions with three response options, the chance level was .33, for questions with two options the chance level
was .50. Binomial tests indicated that mock-jurors gave correct answers to nine out of ten questions at a level significantly higher than chance (see Table 6). Thus, mock-jurors who received the judicial instructions demonstrated near perfect understanding of the information provided to them.
Table 6

Questions Assessing the Comprehension of the Judicial Instructions, and Percentage of Mock-jurors who Gave the Correct Answer

<table>
<thead>
<tr>
<th>Question</th>
<th>Correct answer</th>
<th>Percentage correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>How many stages of memory are involved in the process of making a positive identification?</td>
<td>Three</td>
<td>91%***</td>
</tr>
<tr>
<td>What are the stages of memory that are involved in the process of making a positive identification?</td>
<td>Observation; Retention; Identification</td>
<td>89%***</td>
</tr>
<tr>
<td>Identification evidence in potentially unreliable.</td>
<td>True</td>
<td>98%***</td>
</tr>
<tr>
<td>A witness may think that s/he is accurate and still be mistaken</td>
<td>True</td>
<td>100%***</td>
</tr>
<tr>
<td>What are the three factors that were discussed in the instructions that are relevant to assessing the reliability of a witness’ identification?</td>
<td>Circumstances of observation; the characteristics of the witness; factors concerning the identification</td>
<td>71%</td>
</tr>
<tr>
<td>Which of these factors should not be considered when evaluating the circumstances of observation</td>
<td>Eyewitness’s mood</td>
<td>73%***</td>
</tr>
<tr>
<td>Which of these factors should not be considered when evaluating the characteristics of the witness</td>
<td>Gender of the eyewitness</td>
<td>87%***</td>
</tr>
<tr>
<td>Stress increases the accuracy of a witness’ identification</td>
<td>False</td>
<td>69%***</td>
</tr>
<tr>
<td>An identification procedure is fair if:</td>
<td>The police officer did not know who the suspect was</td>
<td>83%***</td>
</tr>
<tr>
<td>The degree of similarity between the suspect and other people in the photoboard can influence accuracy of identification</td>
<td>True</td>
<td>93%***</td>
</tr>
</tbody>
</table>

Note. ***p < .001
Discussion

The present study tested the efficacy of judicial instructions in sensitizing mock-jurors to the quality of the identification conditions and to eyewitness accuracy presented on video. Contrary to Hypotheses 2 and 3, but in line with previous studies with US-based instructions, the Victorian judicial instructions did not improve mock-jurors’ sensitivity to suggestive identification procedures or to the accuracy of the eyewitness’s identification. Instead, judicial instructions lead to confusion. When mock-jurors viewed the suggestive procedure, those who read the instructions estimated a higher likelihood that the witness correctly identified the culprit compared to those in the control condition. Thus, not only did the judicial instructions not sensitize mock-jurors to the quality of the identification procedure, they exaggerated mock-jurors’ insensitivity to suggestive procedures. The ineffectiveness of the judicial instructions found in this study replicates the results of previous studies with different versions of judicial instructions (Ramirez et al. 1996: Experiment 1; Pawlenko et al. 2013). Moreover, this is the first study, to our knowledge, that applied Bayesian analysis to judicial instructions on eyewitness evidence. Bayesian analysis compares the evidence in support of the null and the alternative hypotheses, giving more weight to the conclusion that judicial instructions did not improve jurors’ sensitivity to witnessing or identification conditions.

It is yet unclear why judicial instructions may sometimes lead to the confusion. One possible explanation is that jurors may have misunderstood the judicial instructions. This explanation is unlikely in our study, however, because mock-jurors gave the correct answer to nine out of the ten comprehension questions. However, we did not assess how the control group would perform on the comprehension test; thus, it is unclear whether providing the instructions improved mock-jurors’ knowledge or whether all mock-jurors had a good understanding of eyewitness issues before participating in the study.
Another possible explanation is that the jurors had little guidance on how to apply the relevant information to the actual case. We cannot rule out this possible explanation in our study because the Victorian judicial instructions provide minimal information on how to integrate the instructions into their decisions. It is important to note, however, that Bayesian analysis did not support this result, indicating that, despite being statistically significant, this finding is not robust.

The inability of the judicial instructions to sensitize mock-jurors to the quality of the identification procedure is especially troubling because the procedure was overly suggestive (single-blind administration, biased instructions, and biased lineup) and the Victorian judicial instructions contained information about these exact biases. In reference to the biased lineup: “Was the identification process conducted fairly? For example, did the other people in the photoboard look sufficiently similar to the accused?” In reference to the biased instructions and single-blind administration: “Was the witness influenced in any other way to identify the accused—for example, by the behaviour of the police?” This demonstrates that simply mentioning factors that are important to their decision is not enough to improve the accuracy of mock-jurors’ judgments. This may not be particularly surprising given that providing mock-jurors with an explicit explanation of how each factor influences eyewitness memory did not improve mock-jurors’ decisions (Jones et al. 2017).

Perhaps even more troubling is that our mock-jurors appropriately reported that the non-suggestive identification procedure was fairer and that the lineup administrator had less influence on the eyewitness (cf. suggestive). Nonetheless, this knowledge did not translate into more not guilty verdicts or greater disbelief in the suggestive procedure condition. Our findings replicate prior work demonstrating that even when jurors understand the suggestiveness of an identification procedure, they may fail to adjust their judgments accordingly (Devenport et al. 2002). It is worth noting, however, that even though mock-
jurors were somewhat sensitive to the suggestiveness of identification procedures, their ratings of lineup presentation were still above the mid-point of the scale (where 1 indicated suggestive presentation and 10 indicated non-suggestive presentation), demonstrating that they underestimated the magnitude of suggestiveness.

Also in line with the previous research and supporting Hypothesis 1, the present study demonstrated that mock-jurors can be sensitive to eyewitness accuracy when identifications are obtained from non-suggestive identification procedures (Beaudry et al. 2015; Reardon and Fisher 2011). However, as hypothesized, viewing identifications from a suggestive procedure impaired this sensitivity. This impairment was evident in mock-jurors’ verdicts, in their assessment of the accuracy of the witness’s identification, and in their confidence in those decisions. These findings underscore the importance of how police conduct their identification procedures (Zimmerman et al. 2012). Police practices affect not only the reliability of an identification made by an eyewitness, but also how jurors evaluate that evidence (e.g., Beaudry et al. 2015; Douglass et al. 2010; Smalarz and Wells 2014).

Eyewitnesses who identify a suspect under suggestive circumstances may make a faster identification decision (Weber et al. 2004) and may appear more confident at the time of the identification and at the trial (Brewer and Wells 2006). More research is needed to determine how jurors evaluate video-recorded identification procedures, which factors influence these evaluations, and how legal safeguards (e.g., expert testimony) may affect jurors’ perceptions of eyewitness testimony and of video-recorded identification procedures.

In summary, mock-jurors believed accurate eyewitnesses more often than inaccurate ones when the identification procedure was non-suggestive. However, mock-jurors’ sensitivity was impaired when the identification procedure was suggestive. Even though mock-jurors understood the Victorian judicial instructions, the instructions did not improve the accuracy of mock-jurors’ judgments and, worse, produced confusion. The Bayesian
analysis supports the null results from our inferential tests; that is, our evidence suggests that judicial instructions had no effect on mock-jurors’ perception of the evidence or their related decisions.

Limitations

One limitation of this study was that it was conducted online, which may have influenced how much attention mock-jurors paid to the stimuli. Nonetheless, surveys administered online produced results equivalent to those obtained in laboratory-based research (e.g., Brock et al. 2012; Lieberman 2008; Riva et al. 2003). Moreover, we included a series of manipulation checks to ensure that we did not include responses from inattentive mock-jurors. Moreover, students participated for course credit and knew that they would be need to provide key pieces of information (i.e., identifying the independent variables, dependent variables, and purpose of the research) to the Research Experience Program team to earn their course credit. Thus, the online nature of the study likely had minimal impact on our results.

A second limitation is that our sample consisted of students rather than real or potential jurors from a community sample. It is worth noting, however, that jury simulation studies found similar results with student and community samples (Jones et al. 2017; Bornstein et al. 2017). Third, mock-jurors did not have the opportunity to deliberate like they would have in a real trial. Although deliberation can increase or decrease individual bias, depending on the circumstances of the case (Devine 2012), it is not yet clear how it would affect their assessment of video-recorded eyewitness evidence. Fourth, the study did not present mock-jurors with a full trial. Although this decreases the ecological validity of the study, there is little reason to suspect that mock-jurors’ performance would increase in real trials. Real trials are much longer and present jurors with more information, which may impair jurors’ attention span and decrease their ability to process information.
**Implications for Future Research**

The results of this study underscore the need to revise the traditional approach to instructing the jury. Years of research into the effectiveness of judicial instructions has demonstrated that even though some changes, such as rewriting the instructions, or changing the time of their presentation, may improve jurors’ comprehension, it appears that none of these minor alterations significantly improve the accuracy of jurors’ judgments (e.g., Ogloff and Rose 2005). Novel methods such as question trails or methods that combine PowerPoint slides with a specifically designed tool for assessing eyewitness evidence (e.g., Safer et al. 2016) may be more effective than simply reading the instructions to the jury. Moreover, it is yet unknown how various legal safeguards affect jurors’ perceptions of video-recorded identification procedures. It is clear, however, that suggestive identification procedures may corrupt eyewitness evidence and bias jurors’ perception of it; thus, omitting suggestive eyewitness evidence from court may be the best solution currently available. In order for this to be an effective strategy, attorneys and judges need to be aware of the damaging effects of suggestive procedures, need to be able to recognize such procedures (e.g., Stinson et al. 1997), and prevent unreliable eyewitness evidence from appearing before a jury.
Acknowledgments

This work was partially supported by the Swinburne University of Technology Postgraduate Research Award. The authors would like to thank Bonnie Kirkman, Jamie Smith-Morvell, Meg Blackie, Molly McQueen, Olivia Bradfield, Roy Groncki, and Travis Edmonds for their thorough reviews and comments on the early version of this manuscript.
Chapter 5. Paper 2

The Effects of Expert Testimony on Mock-Jurors’ Perceptions of Video-Recorded Eyewitness Identification Evidence

Alena Skalon, Jennifer L. Beaudry, & Travis Edmonds

Swinburne University of Technology

Author note: Alena Skalon, Department of Psychological Sciences; School of Health Sciences; Faculty of Health, Arts and Design; Swinburne University of Technology, Hawthorn, Victoria, Australia; email: askalon@swin.edu.au. Jennifer L. Beaudry, Department of Psychological Sciences; School of Health Sciences; Faculty of Health, Arts and Design; Swinburne University of Technology, Hawthorn, Victoria, Australia; email: jbeaudry@swin.edu.au. Travis Edmonds, Department of Psychological Sciences; School of Health Sciences; Faculty of Health, Arts and Design; Swinburne University of Technology, Hawthorn, Victoria, Australia; email: edmondstrav@gmail.com.

Please note that I have made small alterations to this Paper for thesis presentation. These changes were limited to placement of tables and figures, numbering of tables and figures, removing all reference lists (which are compiled at the end of the thesis), and altering some in-text citations to align with this compiled list. As no other changes were made, the referencing style and language requested by Law and Human Behavior were retained.
Abstract

The ideal outcome from providing eyewitness expert testimony in a trial is that it will sensitize jurors to the reliability and accuracy of eyewitness evidence. We present two studies examining the effectiveness of expert testimony when combined with video-recorded genuine eyewitness identification evidence. In Study 1, mock-jurors \((N = 561)\) saw a video-recorded identification procedure and eyewitness testimony that varied in terms of identification accuracy and the suggestiveness of the identification procedures. Mock-jurors viewed traditional expert testimony, expert testimony with visual aids, or no expert testimony (control). In Study 2, we also manipulated whether mock-jurors \((N = 478)\) saw both the eyewitness identification and testimony (as in Study 1) or the eyewitness testimony alone. In both studies, mock-jurors were sensitive to identification accuracy when identification procedures were non-suggestive, but not when they were suggestive. Expert testimony did not affect verdicts. Expert testimony with visual aids did not outperform traditional expert testimony. Study 2 demonstrated that mock-jurors were more willing to believe that the eyewitness correctly identified the culprit when they watched the eyewitness testimony alone compared to the identification and eyewitness testimony combined. In both studies, expert testimony improved mock-jurors’ knowledge of eyewitness factors, but it did not improve their ability to differentiate between correct and false identifications or attenuate the influence of viewing suggestive procedures. Thus, even though mock-jurors understood the content of expert testimony, this understanding did not translate into better decisions.

Keywords: expert testimony, eyewitness evidence, legal safeguards, video-recorded identification procedure.
Public Significance Statement

Expert testimony is a legal safeguard that should increase jurors’ understanding of eyewitness identification issues and improve their sensitivity to the reliability of the evidence, such that they believe reliable evidence and dismiss unreliable evidence. These studies demonstrated that expert testimony did not affect mock-jurors’ verdicts. Mock-jurors’ differentiated between correct and false identifications when identification procedures were non-suggestive, however, sensitivity was lost when identification procedures were suggestive. This highlights the importance of using best practice procedures when obtaining identification evidence from eyewitnesses.
The Effects of Expert Testimony on Mock-Jurors’ Perceptions of Video-Recorded Eyewitness Identification Evidence

Cases of wrongful convictions demonstrate that eyewitness misidentifications contribute to more than 70% of cases overturned through DNA testing (Innocence Project, 2019). This indicates not only the fallibility of human memory but also that triers of fact can be persuaded by this type of evidence, even when eyewitnesses are mistaken. One of the possible explanations for wrongful convictions in cases that include an eyewitness identification is that jurors might be unaware of how various factors can affect identification accuracy (Benton, Ross, Bradshaw, Thomas, & Bradshaw, 2006; Desmarais & Read, 2011). In particular, jurors tend to underestimate the importance of system variables (Shaw, Garcia, & McClure, 1999; Wells, 1978) and, even when they recognized that a procedure was suggestive, that awareness did not influence their final verdict (Devenport, Stinson, Cutler, & Kravitz, 2002). Instead, jurors tend to believe confident eyewitnesses, even when the eyewitness’s reported confidence was artificially inflated (Boyce, Beaudry, & Lindsay, 2007; Douglass & Jones, 2013).

One suggestive identification procedure that can inflate eyewitness confidence is confirming post-identification feedback (e.g., “Good, you identified the suspect”; Wells & Bradfield, 1998; Douglass & Steblay, 2006). In addition to reporting higher confidence, eyewitnesses who received confirming post-identification feedback claimed to have had a better view of the culprit and paid more attention during the crime, along with other factors that jurors use to determine the reliability of eyewitness evidence (see Bradfield & Wells, 2000; Wells & Quinlivan, 2009). Consequently, mock-jurors were more likely to believe eyewitnesses who received confirming post-identification feedback compared to those who did not receive any feedback (Douglass, Neuschatz, Imrich, & Wilkinson, 2010; Smalarz &
Wells, 2014; Beaudry et al., 2015). To date, it is unclear whether legal safeguards can correct for jurors’ overbelief of eyewitnesses who received confirming post-identification feedback.

**Expert Testimony**

People’s lack of knowledge of factors that influence eyewitnesses, coupled with jurors’ tendency to believe eyewitnesses, create a dangerous situation for innocent defendants. One possible way to educate jurors about eyewitness factors is to have an expert witness provide information about the psychology of memory. Unlike most types of expert testimony (e.g., a fingerprint expert), eyewitness experts cannot comment on the issues of a particular case or on a specific eyewitness; rather, they give a general overview of the factors that can influence the reliability of eyewitness evidence.

Expert testimony can have one of four effects on jurors’ decisions: (a) a null effect, where jurors’ perceptions and verdicts are not significantly affected by the expert testimony; (b) confusion, where jurors make decisions that conflict with the content of expert testimony; (c) skepticism, where jurors disregard the eyewitness evidence, regardless of the strength of the evidence; and (d) sensitivity to eyewitness evidence, which can take one of two forms: sensitivity to eyewitness accuracy and sensitivity to witnessing (e.g., lighting, distance) and/or identification conditions (e.g., lineup construction, lineup instructions). In the real world, we want jurors to demonstrate sensitivity to eyewitness accuracy; we want jurors to believe correct identifications and disregard false identifications. To examine jurors’ sensitivity to accuracy, researchers utilize a genuine eyewitness paradigm (Wells, Lindsay, & Ferguson, 1979). In this paradigm, mock-jurors watch the testimony of an eyewitness who saw a mock-crime and made a genuine attempt to identify the culprit. This is the only paradigm that controls ground truth—whether the witness made a correct or false identification. In contrast, research that investigated only sensitivity to witnessing and/or
identification conditions used simulated eyewitness evidence by employing actors who played the role of eyewitnesses.

To date, only two peer-reviewed studies examined whether expert testimony helps jurors to discriminate between eyewitnesses who made correct or false identifications (Wells, Lindsay, & Tousignant, 1980; Martire & Kemp, 2009). In Wells et al.’s (1980) study mock-jurors showed skepticism, in contrast, Martire and Kemp (2009) did not find any significant effect of expert testimony on mock-jurors’ judgments. In terms of studies examining jurors’ sensitivity to witnessing and identification conditions, expert testimony would be considered effective if, after hearing it, jurors believed eyewitnesses who reported good (cf. poor) witnessing conditions and/or who made an identification from a non-suggestive (cf. suggestive) identification procedure. There is some indication that expert testimony can improve jurors’ sensitivity to witnessing and identification conditions (Cutler, Penrod, & Dexter, 1989; Geiselman et al., 2002; Devenport et al., 2002). However, the most common effect of expert testimony is skepticism (e.g., Leippe, Eisenstadt, Rauch, & Seib, 2004; Lindsay, 1994; Jones, Bergold, Dillon, & Penrod, 2017).

To date, the results of eyewitness expert testimony studies portray a rather pessimistic picture. The tendency for expert testimony to produce skepticism is problematic because jurors who disbelieve eyewitnesses are arguably as damaging to the fairness of the criminal justice system as jurors who always believe eyewitnesses. Two findings are especially concerning. First, even when there was some evidence that expert testimony sensitizes jurors to witnessing and identification conditions, it did not translate into verdicts. Second, neither of the studies that used a genuine eyewitness paradigm showed evidence of sensitization to eyewitness accuracy (i.e., Wells et al., 1980; Martire & Kemp, 2009). Moreover, neither of these studies manipulated identification conditions—thus, we do not know whether expert testimony will affect mock-jurors’ sensitivity to witnessing and/or identification conditions in
a genuine eyewitness paradigm. Overall, eyewitness expert testimony cannot be considered an effective trial safeguard because it repeatedly failed to sensitize jurors to the quality of eyewitness evidence (in terms of either witnessing and identification conditions or eyewitness accuracy). Given that the other commonly-used safeguard—judicial instructions—is also largely ineffective (e.g., Martire & Kemp, 2009; Jones et al., 2017), we contend that we must not abandon this legal safeguard, but rather investigate conditions under which expert testimony might be effective.

**Presentation Method of Expert Testimony**

Expert testimony on the psychology of eyewitness identification might potentially be improved by illustrating the key concepts using figures or highlighting specific details using slides with bullet points. Research in other domains (e.g., racial discrimination, quality of scientific evidence) demonstrated that accompanying expert testimony with visual aids improved mock-jurors’ recall of the evidence (Park & Feigenson, 2013) and led to sensitivity that appropriately influenced their verdicts (Jones & Kovera, 2015). To date, one unpublished study evaluated the role of visual aids in expert testimony on eyewitness identification (Binder, 2006, as cited in Park & Feigenson, 2013). Binder presented PowerPoint slides that highlighted key information presented by the expert in text on the slides. She found that these slides improved the perceived credibility of the expert, but did not affect verdicts. Although Binder included visual aids, she did not present any new or supporting information (such as graphs displaying research findings), which may explain the lack of an effect. Given that visual aids have increased the memorability and influence of expert testimony in other fields, more research is needed to explore whether including visual aids can improve mock-jurors’ sensitivity to eyewitness evidence.

**Video-recorded Eyewitness Identification Evidence**
Traditionally, fact-finders have relied on an eyewitness’s testimony and the testimony of the police officer to determine the eyewitness’s confidence and the fairness of the identification procedure. With technological advancements, it is possible to video-record the identification procedure and, thus, provide an objective account of what happened during the identification (Kassin, 1998; Reardon & Fisher, 2011; Wilford & Wells, 2013). A video-recorded identification procedure depicts the behaviour of the eyewitness, including the eyewitness’ confidence at the time of the identification, how quickly the eyewitness made the decision, and how the identification procedure was conducted (Brewer & Wells, 2006; Kaminski & Sporer, 2017; Wixted & Wells, 2017).

There is some evidence that viewing the video-recorded identification procedure improves mock-jurors’ sensitivity to eyewitness accuracy (Reardon & Fisher, 2011); however, Beaudry et al. (2015) identified potential boundary conditions to the benefits of video-recorded identification procedures. In their study, mock-jurors believed correct identifications more than false identifications only when they watched a video of an identification made from a simultaneous, but not sequential, lineup. Furthermore, ratings of belief were higher for eyewitnesses who received feedback than for those who did not receive feedback—regardless of eyewitness accuracy or evidence type. Thus, exposure to the video-recorded identification procedure did not sensitize mock-jurors to the damaging influence of post-identification feedback. One potential explanation for this lack of sensitivity to the identification conditions captured on video is that mock-jurors were not given any information to assist them with evaluating the eyewitness evidence. Thus, research is needed to better understand how expert testimony will influence mock-jurors’ perceptions of video-recorded identification procedures from genuine eyewitnesses.

**Present Studies**

This paper presents two studies that tested the effectiveness of expert testimony in
sensitizing jurors to suggestive identification procedures and to eyewitness accuracy. The first study investigated whether expert testimony (either traditional or enhanced with visual aids) affected mock-jurors’ sensitivity to identification accuracy and identification procedures when mock-jurors viewed both an eyewitness’s video-recorded identification and testimony. The second study examined if mock-jurors’ sensitivity varied according to the type of evidence (same as in Study 1 or eyewitness testimony alone). Both studies were registered prior to the analysis of the full data set, but not before data collection began (https://tinyurl.com/y9kvbv5m).

**Study 1**

In this study, mock-juror participants viewed a mock-trial in which they watched the video-recorded identification procedure and testimony of a genuine eyewitness. The eyewitness had either made a correct or a false identification from a non-suggestive or suggestive identification procedure. The mock-jurors viewed traditional expert testimony, expert testimony accompanied by visual aids, or completed a filler activity (control condition). The key dependent variables of interest were the verdict and mock-jurors’ perceptions of the likelihood that the eyewitness correctly identified the culprit.

We had three key hypotheses for this study. First, we hypothesized an Expert Testimony × Identification [ID] Procedure interaction: expert testimony will sensitize mock-jurors to the quality of ID conditions (hypothesis 1). When viewing an ID obtained from a suggestive procedure, mock-jurors who watch expert testimony with visual aids will be less likely to render a guilty verdict, less likely to believe that the eyewitness correctly identified the culprit, and will rate the administrator as having more influence on the eyewitness compared to the traditional expert testimony, which in turn will be greater than the control condition. We expect no significant differences in guilty verdicts across the expert testimony conditions when mock-jurors viewed IDs obtained from non-suggestive procedures.
We also hypothesized a significant ID Accuracy × ID Procedure interaction: when viewing non-suggestive ID procedures, mock-jurors will be sensitive to ID accuracy, rendering more guilty verdicts and estimating higher likelihood the eyewitness correctly identified the culprit if they viewed a correct ID (cf. false ID; hypothesis 2). When viewing suggestive ID procedures, mock-jurors will believe the evidence regardless of the eyewitness’s accuracy. Finally, we predicted that mock-jurors in expert testimony conditions will demonstrate better knowledge of eyewitness factors compared to mock-jurors in the control condition (hypothesis 3).

**Method**

**Disclosures.** The data for the study (both raw and clean) are available on the Open Science Framework (OSF; https://tinyurl.com/y9kvbv5m). We report all data exclusions and manipulations in the study. All measures and the details regarding how we determined our sample size are available at our OSF project. This project has been approved by Swinburne’s Human Research Ethics Committee (SUHREC) in line with the Australian National Statement on Ethical Conduct in Human Research.

**Design.** This study used a 2 (ID Procedure: suggestive vs. non-suggestive) × 2 (ID Accuracy: correct vs. false) × 3 (Expert Testimony: traditional expert testimony, expert testimony with visual aids, control) between-subjects factorial design.

**Participants.** The final sample included 561 students enrolled in first-year psychology and criminology units (78.40% females, 21.60% males) who participated in exchange for extra credit. Participants’ average age was 33.15 years old (SD = 11.17), and the majority of participants self-identified as Caucasian (76.60%), followed by “other” (11.20%), Asian (10.70%), and Indigenous (1.40%).

**Materials.** This section describes the eyewitness evidence stimuli, the expert testimony, the filler activity used as a control manipulation, and the key dependent variables.
The full questionnaire and transcripts of case summary, attorneys’ arguments, and expert testimony are available on our OSF page (https://tinyurl.com/y9kvbv5m).

**Eyewitness identification and eyewitness testimony stimuli.** We used the eyewitness identification and testimony stimuli from Beaudry et al. (2015). We used stimuli from 16 eyewitnesses (4 eyewitnesses for each of the 4 identification conditions) for stimulus sampling purposes (Wells & Windschitl, 1999), but each participant viewed stimuli from only one eyewitness. During the creation of the stimuli, the eyewitness-participants watched a mock-theft video (theft from an unattended purse); after a brief delay they were asked to participate in an identification procedure and to testify about the crime they viewed and their identification decision. The identification attempt was video-recorded with the lineup administrator and the eyewitness-participant sitting at the table, facing each other. Both parties could be seen in profile. Eyewitness-participants were shown a six-person photographic simultaneous lineup. In the non-suggestive ID procedure condition, the lineup administrator did not know the identity of the suspect and did not comment on the eyewitness’s choice. In the suggestive ID procedure condition, the administrator knew who the suspect was and provided post-identification feedback after the eyewitness made an identification (i.e., “Good, you got the guy!”). Correct eyewitnesses identified the target from a target-present lineup, whereas eyewitnesses who made a false identification identified an innocent lineup member from a target-absent lineup.

After making their identification decision, eyewitnesses reported confidence in their decision in their own words. After a brief delay, the eyewitness-participants provided testimony where they answered questions about the culprit and the identification procedure. The testimony was video-recorded with the eyewitness facing the camera; the person asking the testimony questions was off-camera. The average duration of the identification procedure and the testimony videos combined was 6 min 28 s.
Expert testimony stimuli. We video-recorded an actor playing the role of the expert who, reading from a script, testified in response to examination by the (off-camera) defense attorney. The expert provided general information about the psychology of memory, including the three memory stages—encoding, retention, and retrieval. Then, the expert talked about some of the misconceptions about memory (e.g., memory like a video-recording device) and described how eyewitness research is conducted before delving into a discussion of various factors. The expert explained that eyewitness factors can be separated into three categories, based on whether they occur during encoding, retention, or retrieval. Factors at the encoding phase included the length of exposure to the culprit’s face, lighting conditions, distance, amount of stress experienced by the eyewitness, and the presence of a weapon. Factors that might be present at the retention phase were the delay between the crime and the identification, and discussions with other eyewitnesses. Finally, information about factors at the retrieval stage included lineup instructions, lineup type, and post-identification feedback. The content of the expert testimony was not tailored to the case at hand (e.g., Devenport et al., 2002). The expert testimony with visual aids (e.g., Park & Feigenson, 2013) contained the same information, plus additional illustrations depicted on PowerPoint slides (see Figure 9). For example, mock-jurors were shown graphs illustrating the influence of suggestive instructions, the difference between simultaneous and sequential lineups, and slides summarizing factors that might be present at each phase. The duration of the expert testimony videos was 22 min 48 s; the duration was the same for the traditional expert testimony and for the testimony that included visual aids.
**Figure 9.** Two sample slides that accompanied the expert testimony in the expert testimony with visual aids condition.

**Control.** Mock-jurors in the control condition were asked to memorize random letters and numbers (e.g., TFDVQ) and recall them after a 20-second delay (similar to the task used in Martire & Kemp, 2009). This task took approximately two minutes to complete.

**Judge’s remarks and attorneys’ arguments.** Actors were given scripts and assigned to the role of judge, prosecuting attorney, and defense attorney. Combined, the videos of the judge’s case summary, attorneys’ arguments, and the judge’s closing remarks lasted 4 min 15 s. In the case summary, the judge listed the facts of the case; the opening and closing statements by attorneys presented the facts of the case highlighting either strengths (prosecutor) or weaknesses (defense attorney) of the case.
**Questionnaire.** Our main dependent variables included: verdict (binary: guilty vs. not guilty); confidence-in-verdict, ranging from 0% (*not at all confident*) to 100% (*extremely confident*); and the likelihood that the person identified by the eyewitness was the culprit [ID likelihood], ranging from 0% (*very unlikely*) to 100% (*very likely*). We also created three combined measures: perceptions of the eyewitness (ratings of the witness’s view, attention, and memory; $rs > .57$, $ps < .001$), perceptions of the administrator (ratings of the administrator’s influence, knowledge of the suspect’s identity, and investment in the outcome of the lineup; $rs > .60$, $ps < .001$), and perceptions of the expert (ratings of the expert’s credibility and the usefulness of expert’s testimony, as well as a question asking how much mock-jurors relied on the information provided by the expert; $rs > .57$, $ps < .001$). In addition, mock-jurors indicated whether they noticed the administrator make any comments to the eyewitness after their identification decision [feedback awareness]. The perceptions of the expert and the feedback awareness analyses are available in the supplemental materials.

**Knowledge of eyewitness factors.** Eleven questions assessed mock-jurors’ knowledge of eyewitness factors; five were true/false questions and six were multiple-choice questions.

**Manipulation checks.** Mock-jurors were asked to report whether they watched all of the videos and whether they recognized anyone from the videos.

**Procedure.** Mock-jurors who signed up to participate received a link to the online study (Qualtrics, Provo, UT). Following the link, they read the consent information statement and provided their consent. This study used a full-trial paradigm. Mock-jurors watched a case summary and attorneys’ opening arguments, then they watched a video-recorded identification procedure and the testimony of an eyewitness. Following this, mock-jurors in both expert testimony conditions watched the relevant expert testimony stimuli (traditional or with visual aids). Mock-jurors in the control condition completed the control activity. At the
end of the study, all mock-jurors watched the attorneys’ closing arguments and the judge’s closing remarks before completing the questionnaire.

**Results**

**Data preparation.** Out of the 797 people who began this study, we excluded the data from 70 people who did not complete the study. We excluded three more people who were not assigned a condition due to a programming error. We also excluded participants (n = 62) who took less than 18 minutes (control condition) or less than 39 minutes (expert testimony conditions) to complete the study. We determined these minimum durations based on the combined duration of all videos in the condition. Likewise, we excluded 94 participants who were identified as extremes based on the stem-and-leaf diagram (i.e., took longer than 72 minutes in the control condition or longer than 122 minutes in the expert testimony condition to complete the study). Finally, we excluded four people who reported that they knew at least one of the actors and three more people who indicated that they did not see any videos. The final sample consisted of 561 participants.

**Confidence-in-verdict measure.** Mock-jurors were asked to indicate how confident they were in their verdict decision. We created a confidence-in-verdict measure by multiplying the confidence rating by -1 if they gave a not-guilty verdict and by +1 if they gave a guilty verdict (Tenney et al., 2007). Thus, a confidence-in-verdict rating score of -100% would indicate that a participant rendered a not-guilty verdict and was 100% confident in that verdict; a score of 100% would indicate that a participant rendered a guilty verdict and was 100% confident in that verdict.

**Data analytic approach.** We collapsed data across the four eyewitnesses presented in each of the conditions (i.e., four non-suggestive false identifications, four non-suggestive correct identifications, four suggestive false identifications, four suggestive correct identifications). We analyzed the verdict measure using a binary logistic regression and we
report odds ratios and percentages for this measure. We used ANOVAs to examine the independent and interactive effects of ID accuracy, ID procedure, and expert testimony on the confidence-in-verdict, ID likelihood, perceptions of the eyewitness, perceptions of the administrator, and perceptions of the expert; for these continuous measures we report means and standard deviations (see Table 7). We report 95% confidence intervals for our effect sizes in square brackets. To examine the effect of expert testimony on mock-jurors’ knowledge of eyewitness factors, we used chi-square analyses. We pre-registered that we will use MANOVA to examine whether expert testimony improves mock-jurors’ knowledge; however, we had to change this to chi-square analyses due to the fact that knowledge questions are categorical.

For our continuous measures, we also estimated the Bayes factor (JASP Team, 2018). Bayesian analyses compare the evidence in favor of both the null and the alternative hypotheses and quantifies the evidence for and against these hypotheses (e.g., Wagenmakers et al., 2018; Jarosz & Wiley, 2014). Raftery (1995) suggests the following interpretation of the strength of Bayes factors. An inverse of Bayes factor (BF10) of 1-3 provides weak support for an alternative hypothesis, a Bayes factor of 20-150 provides strong support for an alternative hypothesis, and a Bayes factor bigger than 150 provides very strong support for an alternative hypothesis. JASP does not support Bayesian logistic regression, thus, we did not analyze the verdict measure using Bayesian approach.

Verdict. We analyzed verdict using binary logistic regression, with the main effects and interactions of three independent variables entered hierarchically: ID procedure (the first step), ID accuracy (the second step), expert testimony (third step), ID Procedure × ID Accuracy (the fourth step), ID Procedure × Expert Testimony (the fifth step), ID Accuracy × Expert Testimony (the sixth step), and ID Accuracy × ID Procedure × Expert Testimony (the seventh step). The last three steps did not significantly improve the model, ps > .28, and were
removed. The main effect of expert testimony did not improve the model fit, \( p > .38 \), and was removed. Thus, contrary to Hypothesis 1, expert testimony did not sensitize mock-jurors (at least as measured by verdict) to the suggestiveness of the identification conditions.

The final model with ID procedure, ID accuracy, and the ID Accuracy \( \times \) ID Procedure interaction was statistically significant, \( \chi^2(3) = 11.96, p = .008 \), explained 28\% (Nagelkerke \( R^2 \)) of the variance in verdict, and correctly classified 56.90\% of cases. As predicted in hypothesis 2, the ID Accuracy \( \times \) ID Procedure interaction was significant (see Figure 10), \( B = 2.72, SE = .34, \) Wald \( (1, n = 561) = 8.58, p = .003 \). Of the mock-jurors who watched non-suggestive ID procedures, more mock-jurors who viewed a correct ID rendered a guilty verdict (61.20\%) compared to those who viewed a false ID (41.60\%), \( \chi^2(1) = 10.93, p < .001, \) OR = 2.22 [1.38, 3.56]. This sensitivity to eyewitness accuracy was not present for mock-jurors who viewed IDs obtained from suggestive procedures (correct IDs: 51.10\%; false IDs: 56.30\%), \( \chi^2(1) = 0.73, p = .393, \) OR = 0.81 [0.51, 1.31]. Further investigation demonstrated that the impaired discrimination was driven by an increase in guilty verdicts for false IDs in the suggestive (cf. non-suggestive) condition, \( \chi^2(1) = 6.02, p = .014, \) OR = 1.80 [1.12, 2.90]. The suggestiveness of ID procedures did not significantly affect guilty verdicts for correct IDs, \( \chi^2(1) = 2.90, p = .089, \) OR = 0.66 [0.41, 1.07].

![Bar chart showing percentage of guilty verdicts for correct and false IDs in non-suggestive and suggestive identification procedures. The chart includes error bars and asterisks indicating statistical significance. ](chart.png)
Figure 10. The two-way interaction between identification accuracy and identification procedure on mock-jurors’ verdicts. Error bars reflect 95% confidence intervals.

***p < .001. *p < .05.

Confidence-in-verdict. The ANOVA indicated that there was a main effect of ID accuracy, $F(1, 549) = 4.14, p = .042, \eta^2_p = .01$. Mock-jurors who viewed correct IDs were more confident in their guilty verdict than those who viewed false IDs, $d = 0.17 [-5.58, 5.91]$. However, this main effect should be interpreted within a significant ID Procedure × ID Accuracy interaction, $F(1, 549) = 6.56, p = .011, \eta^2_p = .01$. The confidence-in-verdict measure mimics the results with the binary verdict measure: mock-jurors who watched a non-suggestive ID procedure were sensitive to ID accuracy, but this sensitivity was eliminated for mock-jurors who viewed IDs obtained from a suggestive procedure. Specifically, in the non-suggestive ID condition, mock-jurors who viewed correct IDs were more confident in their guilty verdict, whereas those who viewed false IDs were more confident in their not-guilty verdicts, $t(282) = 3.44, p = .001, d = 0.41 [-7.50, 8.32]$. However, in the suggestive ID condition, mock-jurors reported similar levels of confidence in their verdicts, regardless of the actual accuracy of the ID, $t(275) = 0.57, p = .567, d = 0.07 [-8.15, 8.29]$. All other main effects and interactions were not statistically significant, $Fs < 1.15, ps > .316, \eta^2_p < 0.004$. We should note that, as is typical with this type of scale (e.g., Palmer, Button, Barnett, & Brewer, 2014; Tenney et al., 2007), there was significant variability in mock-jurors’ confidence-in-verdict ratings, as evidenced by the large standard deviations and confidence intervals for the effect sizes.
<table>
<thead>
<tr>
<th></th>
<th>Confidence in verdict</th>
<th>Likelihood ID</th>
<th>Perceptions of the administrator</th>
<th>Perceptions of the eyewitness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non-suggestive</td>
<td>Suggestive</td>
<td>Total</td>
<td>Non-suggestive</td>
</tr>
<tr>
<td>Control</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct ID</td>
<td>28.95 (65.41)</td>
<td>8.78 (67.08)</td>
<td>20.76 (66.51)</td>
<td>63.35 (21.32)</td>
</tr>
<tr>
<td>False ID</td>
<td>-5.19 (70.11)</td>
<td>23.58 (72.48)</td>
<td>10.17 (72.48)</td>
<td>52.35 (23.50)</td>
</tr>
<tr>
<td>Total</td>
<td>13.78 (69.35)</td>
<td>17.26 (70.25)</td>
<td>15.42 (69.62)</td>
<td>58.46 (22.88)</td>
</tr>
<tr>
<td>Expert (traditional)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct ID</td>
<td>19.08 (72.27)</td>
<td>2.44 (71.49)</td>
<td>9.34 (71.90)</td>
<td>54.90 (25.92)</td>
</tr>
<tr>
<td>False ID</td>
<td>-10.82 (70.27)</td>
<td>2.53 (73.25)</td>
<td>-4.22 (71.65)</td>
<td>44.23 (21.26)</td>
</tr>
<tr>
<td>Total</td>
<td>3.23 (72.35)</td>
<td>2.48 (71.89)</td>
<td>2.82 (71.90)</td>
<td>49.24 (24.03)</td>
</tr>
<tr>
<td>Expert (visual)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct ID</td>
<td>13.02 (72.79)</td>
<td>13.49 (66.24)</td>
<td>13.22 (69.60)</td>
<td>54.83 (23.97)</td>
</tr>
<tr>
<td>False ID</td>
<td>-4.51 (60.97)</td>
<td>7.93 (68.59)</td>
<td>1.78 (64.88)</td>
<td>43.89 (20.70)</td>
</tr>
<tr>
<td>Total</td>
<td>4.54 (67.54)</td>
<td>10.41 (67.20)</td>
<td>7.31 (67.25)</td>
<td>49.54 (23.00)</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct ID</td>
<td>21.13 (69.59)***</td>
<td>7.47 (68.35)</td>
<td>14.64 (69.22)*</td>
<td>58.33 (23.69)***</td>
</tr>
<tr>
<td>False ID</td>
<td>-6.77 (66.86)***</td>
<td>12.30 (71.58)</td>
<td>3.00 (69.85)*</td>
<td>46.96 (22.10)***</td>
</tr>
<tr>
<td>Total</td>
<td>7.67 (69.58)</td>
<td>9.98 (69.96)</td>
<td>8.81 (69.72)</td>
<td>52.85 (23.59)*</td>
</tr>
</tbody>
</table>

Note: Standard deviations are in parentheses.
***p < .001, **p < .01, *p < .05.
**ID likelihood.** This question asked mock-jurors to estimate the likelihood that the person identified by the eyewitness was the culprit; thus, it was a more direct and more sensitive measure of mock-jurors’ belief of the eyewitness evidence. All three main effects were significant as was the ID Procedure × ID Accuracy interaction. The main effect of ID accuracy indicated that mock-jurors who viewed correct IDs provided significantly higher likelihood estimates than those who viewed false IDs, $F(1, 549) = 9.68, p = .002, d = 0.26 [-1.65, 2.16]$. The main effect of ID procedure demonstrated that mock-jurors who viewed IDs from suggestive procedures provided significantly higher likelihood estimates compared to those who viewed IDs obtained from non-suggestive procedures, $F(1, 549) = 5.50, p = .019, d = 0.17 [-1.75, 2.08]$.

The main effect of expert testimony indicated that mock-jurors in the control condition provided higher likelihood estimates than those in either expert testimony condition, $F(2, 549) = 4.69, p = .010, \eta^2_p = .02$. However, a post-hoc analysis using Bonferroni adjusted alpha levels of .017 per test (.05/3) revealed that these differences were not statistically significant ($ps < .021, ds < 0.29$). Like other belief measures, and in support of Hypothesis 2, we found a significant ID Procedure × ID Accuracy interaction, $F(1, 549) = 6.19, p = .013, \eta^2_p = .01$. Mock-jurors demonstrated sensitivity to ID accuracy only when the ID procedures were non-suggestive, $t(282) = 4.17, p < .001, d = 0.50 [-2.16, 3.16]$. Mock-jurors who watched suggestive IDs showed no sensitivity, $t(275) = 0.20, p = .838, d = 0.03 [-2.65, 2.70]$. All other main effects and interactions were non-significant, $Fs < 0.68, ps > .507$.

**Bayesian ANOVA.** A three-way Bayesian ANOVA (JASP Team, 2018) supported the findings of the inferential statistics for the ID likelihood measure. The best model included all three main effects and the ID Procedure × ID Accuracy interaction. Under this model, the
data provided stronger support for the alternative hypothesis than for the null hypothesis (BF$_{10} = 82.04$).

**Perceptions of the administrator.** For this aggregate measure, a high score indicates that mock-jurors reported that the administrator had more influence on the eyewitness. A significant main effect of ID procedure revealed that mock-jurors who viewed suggestive procedures rated the administrator as having more influence compared to those who viewed non-suggestive procedures, $F(1, 548) = 627.61, p < .001, d = 2.12$ [1.93, 2.30]. There was also a significant main effect of ID accuracy: mock-jurors who viewed correct IDs perceived the administrator more positively than those who viewed false IDs, $F(1, 548) = 6.25, p = .013, d = 0.19 [-0.08, 0.46]$, but this was a small effect.

In addition to these main effects, there was a significant ID Procedure × Expert Testimony interaction, $F(2, 548) = 3.96, p = .020, \eta^2_p = .01$. Mock-jurors in the control condition who watched non-suggestive IDs assigned lower ratings of administrator influence than those who watched suggestive IDs, $F(1, 202) = 151.22, p < .001, d = 1.73$ [1.40, 2.06]. This sensitivity was exaggerated in the traditional expert testimony condition, $F(1, 178) = 204.26, p < .001, d = 2.15$ [1.82, 2.47] and exaggerated even further in the expert testimony with visual aids condition, $F(1, 174) = 310.79, p < .001, d = 2.67$ [2.38, 2.97]. It is important to note that even though the differences were exaggerated, the differences between control, expert testimony, and expert testimony with visual aids were not statistically significant for non-suggestive and for suggestive ID procedures, $Fs < 7.25, ps > .038, \eta^2_p < 0.02$ (Bonferroni adjusted alpha levels were .017). All other main effects and interactions were non-significant, $Fs < 1.05, ps > .350$.

**Bayesian ANOVA.** A Bayesian ANOVA provided partial support for the findings of the inferential statistics. The best model contained a main effect of ID procedure and a main effect of ID accuracy, but no main effect of expert testimony, and no interactions. This model
demonstrated strong support for the alternative hypothesis ($BF_{10} = 4.42 \times 10^{89}$). A model that included only main effects of ID procedure, expert testimony, and their interaction performed worse ($BF_{10} = 1.69 \times 10^{88}$), even though it still provided support for the alternative hypothesis.

**Perceptions of the eyewitness.** For this aggregate measure, higher scores reflect more positive perceptions of the eyewitness (better view, memory, and attention). The significant main effects of ID procedure, $F(1, 549) = 17.14, p < .001, d = 0.31 [0.13, 0.49]$, and ID accuracy, $F(1, 549) = 32.22, p < .001, d = 0.48 [0.30, 0.66]$, should be interpreted within their significant interaction, $F(1, 549) = 32.68, p < .001, \eta^2_p = .06$. In the non-suggestive ID condition, mock-jurors who viewed correct IDs perceived the eyewitness more positively compared to those who viewed false IDs, $t(282) = 8.55, p < .001, d = 1.01 [0.78, 1.25]$. However, in suggestive ID condition, there were no significant differences between ratings for correct and false IDs, $t(275) = 0.12, p = .906, d = 0.01 [-0.23, 0.26]$. As expected, this finding mimics those found with the ID likelihood measure. Expert testimony did not affect mock-jurors’ perceptions of the eyewitness either independently or interactively, all $ps > .196$.

**Bayesian ANOVA.** A Bayesian ANOVA supported the inferential statistics’ results. Under the best model, which included the ID Procedure $\times$ ID Accuracy interaction and its main effects, the data were much more likely to occur under the alternative hypothesis than under the null hypothesis ($BF_{10} = 5.36 \times 10^{13}$).

**Knowledge of eyewitness factors.** We recoded all knowledge questions into a binary correct or incorrect choice and ran chi-square tests to determine if there were differences between the two types of expert testimony and between the expert testimony and control conditions. There were no differences between traditional expert testimony and expert testimony with visual aids, all $ps > .126$; thus, we collapsed the two conditions and reported
an average percentage for the combined expert testimony conditions (see Table 8 for the percentage of mock-jurors who answered each question correctly). Expert testimony improved mock-jurors’ knowledge compared to the control condition for 9 out of 11 questions.
Table 8

*Percentage of Correct Answers in Control and Expert Testimony Conditions (Study 1)*

<table>
<thead>
<tr>
<th>Question</th>
<th>Correct answer</th>
<th>Control</th>
<th>Expert testimony (combined)</th>
</tr>
</thead>
<tbody>
<tr>
<td>How many stages of memory are involved in the process of making a positive identification? &lt;select from multiple options&gt;</td>
<td>Three</td>
<td>71%</td>
<td>92%***</td>
</tr>
<tr>
<td>What are the stages of memory that are involved in the process of making a positive identification? &lt;select from multiple options&gt;</td>
<td>Observation; Retention; Identification</td>
<td>79%</td>
<td>99%***</td>
</tr>
<tr>
<td>A witness may think that s/he is accurate and still be mistaken.</td>
<td>True</td>
<td>99%</td>
<td>100%</td>
</tr>
<tr>
<td>Stress increases the accuracy of a witness’ identification.</td>
<td>False</td>
<td>64%</td>
<td>79%***</td>
</tr>
<tr>
<td>Which of these factors should <em>not</em> be considered when evaluating the encoding phase? &lt;select from multiple options&gt;</td>
<td>Eyewitness’s mood</td>
<td>84%</td>
<td>94%***</td>
</tr>
<tr>
<td>Choose the most accurate statement: &lt;select from multiple options&gt;</td>
<td>Eyewitness’s memories can change because of discussions with other people</td>
<td>90%</td>
<td>95%*</td>
</tr>
<tr>
<td>Which of these factors should <em>not</em> be considered when evaluating the identification procedure? &lt;select from multiple options&gt;</td>
<td>Whether the eyewitness was tired when the photoboard was conducted</td>
<td>55%</td>
<td>84%***</td>
</tr>
<tr>
<td>Simultaneous photobords are better [than sequential photobords] because eyewitness can compare faces.</td>
<td>False</td>
<td>34%</td>
<td>83%***</td>
</tr>
<tr>
<td>Post-identification feedback is a common procedure and there is nothing wrong with it.</td>
<td>False</td>
<td>77%</td>
<td>84%*</td>
</tr>
<tr>
<td>Post-identification feedback affects:</td>
<td>Confidence in the identification decision</td>
<td>83%</td>
<td>89%</td>
</tr>
<tr>
<td>Photobords should be banned because they are utterly suggestive.</td>
<td>False</td>
<td>72%</td>
<td>91%***</td>
</tr>
</tbody>
</table>

Note. Significant differences between the control and expert testimony conditions are denoted by asterisks.

***$p < .001$. *$p < .05$
Discussion

The focus of the first study was on exploring the effectiveness of traditional expert testimony and expert testimony with visual aids. Our first hypothesis that expert testimony will sensitize mock-jurors to the presence of suggestive procedures was not supported. In fact, there is some (albeit conflicting) evidence to suggest that the only potential effect of expert testimony on mock-jurors’ judgments was skepticism. We found evidence of skepticism on only one dependent variable—likelihood that the eyewitness made a correct identification. Furthermore, the null results from the Bonferroni-corrected post hoc inferential tests conflict with the Bayesian analysis, which provided strong support for the alternative hypothesis (cf. the null hypothesis). Moreover, contrary to our expectation that expert testimony with visual aids would outperform traditional expert testimony, there were no differences between the two conditions on any measures, thus, the second part of hypothesis 1 was not supported. The only potential effect of expert testimony—skepticism effect—in our study replicated decades of research (e.g., Cutler, Dexter, & Penrod, 1990; Jones et al., 2017; Leippe et al., 2004) and extended it to eyewitness evidence that was presented via a video-recorded identification procedure along with the eyewitness’s testimony.

We found strong and consistent support for our second hypothesis: a significant ID Accuracy × ID Procedure interaction emerged for all key measures. Mock-jurors who viewed non-suggestive identification procedures were sensitive to identification accuracy. Mock-jurors who viewed a non-suggestive correct identification were more likely to render a guilty verdict, estimated higher likelihood that the eyewitness correctly identified the culprit, and perceived the eyewitness to be more reliable compared to those who viewed a non-suggestive false identification. However, mock-jurors who viewed suggestive identification procedures did not discriminate between correct and false identifications. Similarly to Devenport et al. (2002), suggestive identification procedures affected mock-jurors’ belief of eyewitness
evidence even though they were aware of the administrator’s influence. The damaging effect of suggestive identification procedures on mock-jurors’ sensitivity to identification accuracy provides further support to Beaudry et al.’s (2015) findings that presenting mock-jurors with video-recorded suggestive identification procedures impair (rather than assist) sensitivity to identification accuracy. Moreover, our results suggest that courts cannot rely on expert testimony to correct for the damaging effects of suggestive identification procedures on people’s evaluations of identification evidence.

As predicted in our third hypothesis, expert testimony improved mock-jurors’ knowledge for 9 out of 11 questions. For example, mock-jurors who watched expert testimony were more knowledgeable about the effects of stress and the benefits of sequential lineups. Without the expert testimony, mock-jurors already demonstrated very good understanding that an eyewitness can think that he/she is accurate and still be mistaken. Likewise, mock-jurors also demonstrated relatively good understanding that post-identification feedback affects confidence in the identification decision without the aid of expert testimony. Critically, there were no significant differences between the two expert testimony conditions; thus, including visual aids did not improve comprehension above and beyond the traditional expert testimony. Taken together, although mock-jurors understood the content of the expert testimony and demonstrated a greater understanding of eyewitness issues, the legal safeguard did not improve their assessment of the eyewitness evidence compared to those in the control condition.

Our study was the first one, to our knowledge, that combined the presentation of video-recorded identification procedure and expert testimony. Traditionally, researchers presented mock-jurors with only an eyewitness’s testimony (e.g., Cutler, Dexter, & Penrod, 1989; Devenport et al., 2002; Jones et al., 2017), which was the approach used in the two studies that examined the effects of expert testimony using the genuine eyewitness paradigm.
On their binary belief measure, one of these studies found evidence of skepticism (Wells et al., 1980) whereas the other found null effects (Martire & Kemp, 2009). The reason for the conflicting findings is unclear. Our study adds to the conflicting literature because we found a null effect of expert testimony on our binary verdict measure and a potential skepticism effect on the continuous belief measure assessing mock-jurors’ perceptions of the likelihood that the eyewitness correctly identified the culprit. Furthermore, our study combines the eyewitness’ testimony and video-recorded identification decision; thus, it is difficult to determine the effects of expert testimony on mock-jurors’ judgments separate from the effects of watching the eyewitness make their decision. Accordingly, we ran Study 2 to separate the effects of expert testimony from the effects of the video-recorded identification procedure.

**Study 2**

Study 2 investigated whether the inclusion of the video-recorded identification influenced the effectiveness of expert testimony. A video-recorded identification procedure is a fundamentally different piece of evidence than an eyewitness’s testimony. For example, Reardon and Fisher (2011) demonstrated that mock-jurors who watched a video-recorded identification procedure and the eyewitness testimony were better at discriminating between correct and mistaken identifications than mock-jurors who watched the eyewitness testimony alone. When jurors watch video-recorded identification procedure, they are exposed to aspects of the eyewitness’s decision, including initial confidence and decision time as well as verbal and nonverbal behavior that are unlikely to be appropriately conveyed in the eyewitness’s testimony. In essence, they are watching the evidence itself, rather than an eyewitness’s recollection of the evidence. Thus, one potential explanation for consistent null effects of expert testimony in Study 1 is that the video-recorded identification evidence overpowered any potential influence that expert testimony might have had on mock-jurors’ decisions. In addition to addressing this issue, Study 2 also allows us to examine whether
expert testimony can correct for the damaging effects of suggestive identification procedures on people’s sensitivity to identification accuracy when presented via eyewitness testimony (Smalarz & Wells, 2014). We manipulated whether mock-jurors watched an eyewitness’s video-recorded identification and testimony (ID + testimony) or testimony alone. In this study, we manipulated the presence of expert testimony, but removed the manipulation of the type of expert testimony because the findings from Study 1 demonstrated that including visual aids did not outperform the traditional expert testimony. Thus, half of the mock-jurors watched the traditional expert testimony, whereas the other half completed a filler activity.

Our first three hypotheses for this study were similar to those in the first study. Specifically, we hypothesized an Expert Testimony × ID Procedure interaction for our key dependent variables (hypothesis 1; however, unlike in Study 1, we did not examine visual aids). We also predicted a significant ID Accuracy × ID Procedure interaction (hypothesis 2) and expected that expert testimony would improve mock-jurors’ knowledge of eyewitness factors (hypothesis 3). Finally, we had an additional hypothesis for this study. We expected that any effect of expert testimony on mock-jurors’ judgments would emerge only in the testimony condition (cf. ID + testimony). In other words, a significant Expert Testimony × Evidence × ID Procedure interaction should emerge (hypothesis 4). Mock-jurors who watched expert testimony and saw only the eyewitness testimony will be sensitive to the quality of identification conditions (i.e., more guilty verdicts in non-suggestive condition than suggestive condition). On the other hand, mock-jurors in the control condition who watched eyewitness testimony will not be sensitive to identification conditions. As demonstrated in Study 1, we expected that mock-jurors who watched both ID + testimony will not be sensitive to the quality of identification conditions, regardless of the presence of expert testimony. Although we were also interested in how identification accuracy would interact with expert testimony, type of evidence, and suggestiveness of the identification procedure,
we did not plan to examine the four-way interaction given the concerns about higher order interactions and adequate power (Giner-Sorolla, 2018; Simonsohn, 2014).

Method

Disclosures. The data for the study (both raw and clean) are available on the OSF (https://tinyurl.com/y9kvbv5m). We report all data exclusions and manipulations in the study. Please see our OSF project for all measures and the details regarding how we determined our sample size. This research has been approved by Swinburne’s Human Research Ethics Committee (SUHREC) in line with the Australian National Statement on Ethical Conduct in Human Research.

Design. This study used a 2 (ID Procedure: suggestive vs. non-suggestive) × 2 (ID Accuracy: correct vs. false) × 2 (Evidence: video ID + testimony vs. testimony only) × 2 (Expert Testimony: expert testimony vs. control) between-subjects factorial design.

Participants. The final sample included 478 students enrolled in first-year psychology and criminology units (76.40% females, 23.40% males) who participated in the study in exchange for extra course credit. The average age of the participants was 30.96 years old (SD = 10.78), and the majority of participants self-identified as Caucasian (78.50%), followed by “other” (12.80%), Asian (7.30%), and Indigenous (1.50%).

Materials. Study 2 used the same materials and followed the same procedure as Study 1 with three exceptions. First, we manipulated the evidence presented to the mock-jurors: half saw both the ID + testimony (as in Study 1), whereas the other half saw only eyewitness testimony. Second, we removed the expert testimony with visual aids condition because it had no effect in Study 1. Thus, mock-jurors were randomly assigned to one of two expert testimony conditions: half of the mock-jurors watched the traditional expert testimony, whereas the other half completed the filler activity used in Study 1 (control condition). Third, we added two knowledge questions testing mock-jurors’ knowledge of the weapon presence
effect and potential for media reports to affect eyewitnesses’ memories. We also removed a question asking mock-jurors whether photoboards are too suggestive and should not be used.

**Results**

**Data preparation.** Out of the 659 people who began this study, we excluded the data from 43 people who did not complete the study. We excluded participants (n = 46) who took less than 18 minutes (control condition, ID + testimony) or less than 17 minutes (control condition, testimony only), less than 39 minutes (expert testimony, ID + testimony) or less than 37 minutes (expert testimony, testimony only) to complete the study. These minimum durations were determined based on the combined duration of all videos in each condition. Likewise, we excluded 86 people who were identified as extremes based on the stem-and-leaf diagram (i.e., took longer than 74 minutes in the control condition or longer than 109 minutes in the expert testimony condition to complete the study). Finally, we excluded three people who knew at least one of the actors and three more people indicated that they did not see any videos. The final sample consisted of 478 participants.

**Data analytic approach.** We used a similar approach as in Study 1, with modifications to examine our specific hypotheses (e.g., including evidence type).

**Verdict.** We analyzed verdict using binary logistic regression. We entered variables hierarchically: ID accuracy and ID procedure (step 1), Evidence (step 2), Expert Testimony (step 3), ID Accuracy × ID Procedure (step 4) Expert Testimony × Evidence (step 5), Expert Testimony × ID Procedure (step 6), Evidence × ID Procedure (step 7), Expert Testimony × Evidence × ID Procedure (step 8). This full model was not statistically significant, \(\chi^2(9) = 5.45, p = .794\) (see Table 13 for the percentage of mock-jurors who rendered a guilty verdict in each condition). On average, 56.10% of mock-jurors rendered a guilty verdict, the ratings varied from 34.80% to 73.10%.
**Confidence-in-verdict.** Similarly to the results obtained with the binary verdict measure, the ANOVA indicated that there were no significant main effects or interactions (see Table 14), \(Fs < 3.09, ps > .080\).

**ID likelihood.** A significant main effect of ID accuracy indicated that mock-jurors who watched correct IDs were more likely to believe that the eyewitness correctly identified the culprit than mock-jurors who watched false IDs (see Table 9 for descriptive statistics), \(F(1, 462) = 7.52, p = .006, d = 0.28 [-1.74, 2.30]\). A significant main effect of evidence demonstrated that mock-jurors who watched eyewitness testimony alone were more willing to believe that the eyewitness correctly identified the culprit compared to those who watched the ID + testimony, \(F(1, 462) = 8.46, p = .004, d = 0.29 [-1.73, 2.31]\).

These main effects should be interpreted within the significant ID Accuracy × Evidence × Expert Testimony interaction, \(F(1, 462) = 4.52, p = .034, \eta^2_p = .01\). In the control condition, mock-jurors who watched ID + testimony were sensitive to ID accuracy, \(d = 0.39 [-4.10, 4.87]\), whereas mock-jurors who watched testimony-only did not differentiate between correct and false IDs, \(t(120) = 2.07, p = .040, d = 0.01 [-3.73, 3.76]\). The pattern was reversed in the expert testimony condition. Mock-jurors who watched the testimony-only, differentiated between correct and false IDs; however, belief rates were similar for mock-jurors who watched the ID + testimony, \(t(123) = 2.82, p = .006, d = 0.51 [-3.03, 4.05]\). All other main effects and interactions were non-significant, all \(Fs < 2.17, ps > 14\).
Table 9

Mock-jurors’ Mean Ratings for the Likelihood that the Eyewitness Correctly Identified the Culprit as a Function of Expert Testimony, Identification Accuracy, and Type of Evidence (Study 2)

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th></th>
<th>Expert</th>
<th></th>
<th>Overall</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Correct ID</td>
<td>False ID</td>
<td>Correct ID</td>
<td>False ID</td>
<td>Correct ID</td>
<td>False ID</td>
</tr>
<tr>
<td>ID + testimony</td>
<td>58.64 (25.15)*</td>
<td>48.92 (25.69)*</td>
<td>56.26 (20.81)</td>
<td>52.29 (21.31)</td>
<td>57.45 (23.00)</td>
<td>50.28 (23.98)</td>
</tr>
<tr>
<td>Testimony</td>
<td>60.31 (22.15)</td>
<td>60.62 (22.05)</td>
<td>64.59 (20.00)**</td>
<td>54.32 (20.69)**</td>
<td>62.32 (21.20)</td>
<td>57.45 (21.52)</td>
</tr>
<tr>
<td>Total</td>
<td>59.62 (23.35)</td>
<td>54.29 (24.71)</td>
<td>60.90 (20.69)</td>
<td>53.42 (20.89)</td>
<td>60.24 (22.07)**</td>
<td>53.90 (23.01)**</td>
</tr>
</tbody>
</table>

Note. Standard deviations are in parentheses.

***p < .001. **p < .01. *p < .05.
**Bayesian ANOVA.** Contrary to the inferential statistics, a Bayesian ANOVA indicated that a model that included only the main effects of ID accuracy and evidence outperformed all other models ($BF_{10} = 63.90$). In contrast, the model with the ID Accuracy × Expert Testimony × Evidence performed worse ($BF_{10} = 0.04$), indicating a lack of robustness in the three-way interaction.

**Perceptions of the administrator.** The significant main effect of ID procedure indicated that mock-jurors who watched suggestive ID procedures correctly assigned higher ratings of administrator influence ($M = 7.05; SD = 2.24$), compared to mock-jurors who watched non-suggestive ID procedures ($M = 3.14; SD = 2.34$), $F(1, 462) = 345.02, p < .001, d = 1.71 [1.51, 1.92]$. In other words, mock-jurors were sensitive to the suggestiveness of the ID procedures, but the main effect should be interpreted within the significant ID Procedure × ID Accuracy interaction (see Table 10), $F(1, 462) = 5.16, p = .024, \eta^2_p = .01$. When they watched non-suggestive ID procedures, mock-jurors who saw a false ID assigned higher ratings of the administrator influence compared to those who watched a correct ID, $t(242) = 2.78, p = .006, d = 0.36 [0.07, 0.65]$. There was no significant difference when ID procedures were suggestive, $t(232) = 0.49, p = .626$.

The ID Procedure × Expert Testimony also demonstrated that mock-jurors were sensitive to ID procedures in both conditions; however, expert testimony exaggerated the sensitivity, $F(1, 462) = 5.82, p = .016, \eta^2_p = .01$. Mock-jurors in the control condition who watched suggestive ID procedures rated the administrator as having more influence than those who watched non-suggestive ID conditions, $t(252) = 11.70, p < .001, d = 1.49 [1.21, 1.77]$. Expert testimony further sensitized mock-jurors to the suggestiveness of the procedures, $t(222) = 14.69, p < .001, d = 1.98 [1.69, 2.28]$. In addition, there was a significant ID Procedure × Evidence Type interaction, $F(1, 462) = 6.51, p = .011, \eta^2_p = .01$. Mock-jurors in the testimony-only condition assigned higher ratings of administrator influence when ID
procedures were suggestive than non-suggestive, $t(255) = 11.92, p < .001, d = 1.49 [1.21, 1.78]$. This sensitivity was exaggerated for mock-jurors who watched ID + testimony, $t(219) = 14.75, p < .001, d = 2.00 [1.70, 2.29]$. However, the differences between two types of evidence were not statistically significant for either of the ID procedure conditions, $t s < 1.65$, $p s > 100$.

Table 10

*Mock-Jurors’ Mean Ratings for the Perceptions of the Administrator According to the Two-Way Interactions between ID Procedure and Expert Testimony, Type of Evidence, and ID Accuracy (Study 2)*

<table>
<thead>
<tr>
<th>ID procedure</th>
<th>Non-suggestive</th>
<th>Suggestive</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID accuracy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct ID</td>
<td>2.73 (2.40)$_a$</td>
<td>7.13 (2.10)</td>
</tr>
<tr>
<td>False ID</td>
<td>3.56 (2.20)$_a$</td>
<td>6.98 (2.37)</td>
</tr>
<tr>
<td>Expert testimony</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>3.30 (2.34)$_{ac}$</td>
<td>6.74 (2.29)$_a$</td>
</tr>
<tr>
<td>Expert testimony</td>
<td>2.91 (2.32)$_{bc}$</td>
<td>7.33 (2.17)$_b$</td>
</tr>
<tr>
<td>Evidence type</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Testimony</td>
<td>3.37 (2.49)$_a$</td>
<td>6.84 (2.17)$_a$</td>
</tr>
<tr>
<td>ID + testimony</td>
<td>2.89 (2.15)$_b$</td>
<td>7.32 (2.32)$_b$</td>
</tr>
</tbody>
</table>

*Note. Standard deviations are reported in parentheses. Means that share subscripts are significantly different within each two-way interaction (ID Procedure x ID Accuracy; ID Procedure x Expert Testimony; ID Procedure x Evidence Type).*

**Bayesian ANOVA.** Again, contrary to inferential statistics, the best model included only ID Procedure ($BF_{10} = 1.12 \times 10^{55}$). A model that included three main effects and an ID
Procedure × Expert Testimony interaction fared worse ($BF_{10} = 4.59 \times 10^{53}$). Thus, the main effect of ID procedure provides the best explanation for the data.

**Perceptions of the eyewitness.** We found a significant main effect of evidence on mock-jurors’ perceptions of the eyewitness. Mock-jurors who watched the testimony only rated the eyewitness more positively ($M = 5.34; SD = 2.10$) than those who viewed the ID + testimony ($M = 4.88; SD = 2.34$), $F(1, 462) = 5.07, p = .025, d = 0.21 [0.01, 0.41]$. There was also a significant main effect of ID accuracy indicating that mock-jurors who watched correct IDs perceived the eyewitness as more reliable ($M = 5.76; SD = 2.14$), compared to mock-jurors who watched false IDs ($M = 4.52; SD = 2.13$), $F(1, 462) = 39.81, p < .001, d = 0.58 [0.39, 0.77]$. These main effects were qualified by two significant interactions (see Table 11). The ID Accuracy × ID Procedure interaction, $F(1, 462) = 19.14, p < .001, \eta^2_p = .04$, again demonstrated that mock-jurors were sensitive to ID accuracy when they viewed non-suggestive ID procedures, $t(242) = 7.52, p < .001, d = 0.97 [0.71, 1.23]$, but not suggestive ones, $t(232) = 1.55, p = .121, d = 0.20 [-0.07, 0.47]$. The ID Accuracy × Expert Testimony interaction demonstrated that mock-jurors were sensitive to ID accuracy in both conditions, $F(1, 462) = 5.41, p = .020, \eta^2_p = .01$. Mock-jurors in the control condition who watched correct IDs perceived the witness as more reliable than those who watched false IDs, $t(252) = 3.56, p < .001, d = 0.45 [0.18, 0.72]$. The same pattern emerged for expert testimony but with a more exaggerated difference in ratings between correct and false IDs, $t(222) = 5.51, p < .001, d = 0.74 [0.47, 1.01]$. 

143
Table 11

*Mock-Jurors’ Perceptions of the Eyewitness According to the Two-Way Interactions of ID Accuracy with Expert Testimony and ID Procedure (Study 2)*

<table>
<thead>
<tr>
<th>ID accuracy</th>
<th>Correct</th>
<th>False</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID procedure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-suggestive</td>
<td>6.08 (2.23)_{ab}</td>
<td>4.05 (1.97)_{ac}</td>
</tr>
<tr>
<td>Suggestive</td>
<td>5.40 (1.99)_{b}</td>
<td>4.98 (2.19)_{c}</td>
</tr>
<tr>
<td>Expert testimony</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>5.60 (2.16)_{a}</td>
<td>4.62 (2.21)_{a}</td>
</tr>
<tr>
<td>Expert</td>
<td>5.94 (2.11)_{b}</td>
<td>4.41 (2.05)_{b}</td>
</tr>
</tbody>
</table>

*Note.* Standard deviations are reported in parentheses. Means that share subscripts are significantly different within each two-way interaction (ID Accuracy x ID Procedure; ID Accuracy x Expert Testimony).

**Bayesian ANOVA.** A Bayesian ANOVA provided partial support for the findings of the inferential statistics. The model that outperformed all other models included only main effects of ID procedure, ID accuracy, and their interaction ($BF_{10} = 9.79 \times 10^8$). A model that included the ID Accuracy $\times$ Expert Testimony interaction performed much worse ($BF_{10} = 616,696.66$). Thus, the finding that expert testimony exaggerated mock-jurors’ positive perceptions of accurate eyewitnesses and negative perceptions of mistaken eyewitnesses (compared to the control condition) is less robust than the ID Procedure $\times$ ID Accuracy interaction.

**Knowledge of eyewitness factors.** Overall expert testimony improved knowledge for 8 out of 12 questions (see Table 12). This finding suggests that mock-jurors paid attention to the expert testimony and understood its content.
Table 12
Percentage of Correct Answers in Control and Expert Testimony Conditions (Study 2)

<table>
<thead>
<tr>
<th>Question</th>
<th>Correct answer</th>
<th>Control</th>
<th>Expert testimony</th>
</tr>
</thead>
<tbody>
<tr>
<td>How many stages of memory are involved in the process of making a positive identification? &lt;select from multiple options&gt;</td>
<td>Three</td>
<td>76%</td>
<td>93%***</td>
</tr>
<tr>
<td>What are the stages of memory that are involved in the process of making a positive identification? &lt;select from multiple options&gt;</td>
<td>Observation; Retention; Identification</td>
<td>81%</td>
<td>99%***</td>
</tr>
<tr>
<td>A witness may think that s/he is accurate and still be mistaken.</td>
<td>True</td>
<td>99%</td>
<td>100%</td>
</tr>
<tr>
<td>Stress increases the accuracy of a witness’ identification</td>
<td>False</td>
<td>72%</td>
<td>85%***</td>
</tr>
<tr>
<td>Which of these factors should not be considered when evaluating the encoding phase? &lt;select from multiple options&gt;</td>
<td>Eyewitness’s mood</td>
<td>87%</td>
<td>94%*</td>
</tr>
<tr>
<td>Choose the most accurate statement: &lt;select from multiple options&gt;</td>
<td>Eyewitness’s memories can change because of discussions with other people</td>
<td>90%</td>
<td>92%</td>
</tr>
<tr>
<td>Individuals who witness a crime are unlikely to be influenced by a news broadcast of the event*</td>
<td>False</td>
<td>74%</td>
<td>79%</td>
</tr>
<tr>
<td>The presence of a weapon does NOT affect the ability to describe or identify the perpetrator*</td>
<td>False</td>
<td>88%</td>
<td>99%***</td>
</tr>
<tr>
<td>Which of these factors should not be considered when evaluating the identification procedure? &lt;select from multiple options&gt;</td>
<td>Whether the eyewitness was tired when the photoboard was conducted</td>
<td>51%</td>
<td>80%***</td>
</tr>
<tr>
<td>Simultaneous photoboard are better [than sequential photoboard] because eyewitness can compare faces</td>
<td>False</td>
<td>36%</td>
<td>77%***</td>
</tr>
<tr>
<td>Post-identification feedback is a common procedure and there is nothing wrong with it</td>
<td>False</td>
<td>90%</td>
<td>89%</td>
</tr>
<tr>
<td>Post-identification feedback affects: &lt;select from multiple options&gt;</td>
<td>Confidence in the identification decision</td>
<td>76%</td>
<td>87%**</td>
</tr>
</tbody>
</table>

Note. Significant differences between the control and expert testimony conditions are denoted by asterisks.
*Question not present in Study 1.
***p < .001. **p < .01. *p < .05.
Discussion

The second study investigated the effectiveness of traditional expert testimony and tested whether the effectiveness of expert testimony varied as a function of the evidence type. Hypothesis 1—that expert testimony will sensitize mock-jurors to identification procedures—received partial support in this study. Even though this interaction did not emerge for other key dependent variables, mock-jurors who watched expert testimony assigned higher ratings of administrator influence when identification procedures were suggestive, compared to non-suggestive. However, we did not predict that this sensitivity will be present in the control condition.

Hypothesis 2—that mock-jurors will be sensitive to identification accuracy when identification procedures are non-suggestive, but not when they are suggestive—was supported for only one measure. Mock-jurors who viewed non-suggestive identification procedures appropriately perceived eyewitnesses who made a correct identification as more reliable than those who made a false identification. This sensitivity was not present in mock-jurors who watched suggestive procedures. A similar pattern also emerged for the perceptions of the administrator influence. The Bayesian analyses provided support for an ID Procedure × ID Accuracy interaction on the perceptions of eyewitnesses, but not on the perceptions of administrator influence.

Similarly to Study 1, and supporting Hypothesis 3, findings indicated that expert testimony improved mock-jurors’ knowledge of eyewitness factors for 8 out of 12 questions. The four questions that were not different between control and expert testimony conditions assessed mock-jurors’ understanding that the witness can think that he/she is accurate and still be mistaken, that media can affect eyewitness memories, understanding of the circumstances under which the eyewitness’s memories can change, and knowledge of post-identification feedback.
In regards to Hypothesis 4, the type of evidence did not affect mock-jurors’ verdicts. However, a three-way interaction between expert testimony, identification accuracy, and evidence has emerged for mock-jurors’ ratings of the likelihood that the eyewitness made a correct identification. Mock-jurors in the control condition were sensitive to identification accuracy when they watched video identification and testimony (cf. testimony-only), supporting findings of Reardon & Fisher (2011). On the other hand, those who watched expert testimony were sensitive to identification accuracy when they watched eyewitness testimony (cf. ID + testimony). This mirrored pattern of results lends support to the idea that there are fundamental differences between jurors’ perceptions of eyewitness testimony and video-recorded identification procedure. If expert testimony is effective, it may only influence jurors’ when they view the eyewitness’s testimony rather than the video-recorded identification procedure.

It is worth noting, however, that Bayesian analysis did not support this interaction. According to a Bayesian ANOVA, the main effects of identification accuracy and evidence are more likely to provide the best fit. According to this model, mock-jurors who watched a correct identification were more likely to believe that the eyewitness correctly identified the culprit than those who watched a false identification. Mock-jurors who watched eyewitness testimony alone were also more likely to believe the eyewitness than mock-jurors who watched both video identification and eyewitness testimony, regardless of the actual identification accuracy. According to the Bayesian analyses, expert testimony did not affect mock-jurors’ judgments. Thus, future research should seek to replicate this effect and investigate the mechanism behind this interaction.

**General Discussion**

Our study was the first, to our knowledge, to explore the effectiveness of expert testimony in sensitizing mock-jurors to the quality of video-recorded genuine eyewitness
identification evidence. Expert testimony did not improve mock-jurors’ sensitivity to the suggestiveness of identification procedures. Furthermore, there is some evidence that expert testimony produced skepticism—mock-jurors who watched expert testimony (cf. control) provided lower likelihood estimates that the eyewitness correctly identified the culprit, independent of the suggestiveness of the identification procedure or the accuracy of the identification. This provides further support to the conclusion that skepticism is the most common effect of expert testimony (e.g., Cutler, Dexter, & Penrod, 1990; Jones et al., 2017; Leippe et al., 2004). This skepticism effect, however, did not emerge with the less sensitive, but more ecologically valid, binary verdict measure. Despite promising results obtained with enhanced expert testimony in other areas (e.g., Park & Feigenson, 2013; Jones & Kovera, 2015), supplementing expert testimony with visual aids had no effect on mock-jurors’ judgments. Thus, this is the second study, in addition to Wells et al. (1980), using a genuine eyewitness paradigm to demonstrate that expert testimony leads to skepticism.

Expert testimony improved mock-jurors’ knowledge of eyewitness factors. Mock-jurors understood the information provided by the expert, but this knowledge did not translate into more accurate assessments of the evidence. Interestingly, those in the expert testimony condition demonstrated a better understanding that post-identification feedback affects confidence of the eyewitness than the control group in study 2 (76% vs. 87%, respectively), but there was no significant difference between the two groups in study 1 (83% vs. 89%, respectively). We do not have an explanation for this discrepancy.

Study 1 also demonstrated that mock-jurors believed eyewitness evidence that was obtained using suggestive identification procedures more often than the evidence obtained using non-suggestive identification procedures. These results mimic those of Devenport et al. (2002). Even though mock-jurors appropriately rated the suggestive identification procedures as more suggestive than the non-suggestive procedures, this awareness did not translate into
verdicts. The damaging effects of suggestive identification procedures were supported by Bayesian analysis. Even though mock-jurors were sensitive to identification accuracy when identification procedures were non-suggestive, this sensitivity was lost when the identification procedures were suggestive. This further supports Beaudry et al. (2015). It is worth noting, however, that in Study 2 the suggestiveness of identification procedures did not affect mock-jurors’ verdicts or belief that the eyewitness correctly identified the culprit. In addition, the Bayesian factors obtained in Study 2 are much smaller than those obtained in Study 1 and often contradicted inferential statistics results, which might indicate that the findings of Study 2 are less robust.

Taken together, our findings suggest that the reliability of eyewitness identifications might be irreversibly damaged by suggestive identification procedures and that expert testimony repeatedly failed to correct the damage. The lack of a significant effect of visual aids above traditional expert testimony adds to the argument that making changes to the existing legal safeguards may not be sufficient to alert jurors to the role of suggestive identification procedures on eyewitness identifications. Instead, a new method of educating jurors might be more promising—a method that not only informs the jury about eyewitness factors, but also provides an instrument that would help the jury to evaluate the quality of the evidence (e.g., I-I-Eye, Pawlenko, Safer, Wise, & Holfeld, 2013).

Technological developments and the widespread use of video-recording devices are changing the way various types of evidence are presented to juries. For example, with the proliferation of body-worn cameras, jurors might be exposed to both the police officers’ testimony and the relevant video footage (Jones, Crozier, & Strange, 2017). Likewise, an eyewitness’s testimony might be accompanied by a video-record of their identification procedure. We know relatively little about how jurors will interpret video evidence of various kinds and how it will affect their judgments of guilt. Overall, we agree that video-recording
of identification procedures should become routine (Kassin, 1998); however, whether this evidence should be presented to jurors is a separate question. Our study aimed to shed light on this issue and found that mock-jurors were somewhat more skeptical of eyewitness evidence when they watched both the video-recorded identification procedure and eyewitness testimony, compared to testimony alone. This is in line with Beaudry et al. (2015), who found that mock-jurors who watched eyewitness testimony had a bias to believe the identification. Furthermore, Study 2 highlighted that the video-recorded identification procedure and eyewitness testimony are fundamentally different types of evidence. Mock-jurors who watched expert testimony were sensitive to identification accuracy when they watched eyewitness testimony, but not when they watched both video-recorded identification procedure and eyewitness testimony. This is contrary to the pattern of results that emerged for the control group. The criminal justice system needs to embrace that emerging types of evidence might be different from traditional evidence and devise new rules for handling it. Thus, until further research has explored the benefits and risks of presenting video-recorded identification procedures, we contend that we must be cautious about recommending that lawyers should present these videos in court (e.g., Douglass & Jones, 2013).

**Limitations**

Due to the long duration of the videos, mock-jurors may have disengaged with the expert testimony. Although the videos are half an hour long, they are still much shorter than the duration of real trials. Moreover, mock-jurors evidently paid attention to expert testimony as demonstrated by their increased knowledge of eyewitness factors. Another limitation of this study is the difference in the duration of the control condition and the expert testimony condition. We based the filler task in our control condition on previously published work with a similar design (i.e., Martire & Kemp). It is unclear how this difference could have
affected the results; however, future research could use control conditions of comparable duration to the experimental condition to remove this artefact.

Other potential limitations include the lack of deliberation and the use of students as mock-jurors. Little is known about the effects of deliberation in cases that involve video-recorded identification evidence—this should be explored in future research. Jury simulation studies demonstrate no appreciable difference between students and community members (Jones et al., 2017; Bornstein et al., 2017). We would argue that the results of these studies can be extrapolated to jury-eligible community members.

**Future Directions**

Future research should further investigate the effectiveness of various legal safeguards using the genuine eyewitness paradigm. This paradigm is the only type that permits the manipulation of identification accuracy, thus, exposing jurors to the whole spectrum of eyewitnesses’ behavior, including verbal and non-verbal cues that might be indicative of accuracy. In addition, the interaction between the suggestiveness of identification procedures and identification accuracy would not have been possible to examine using simulated eyewitness evidence. Furthermore, there is a lack of knowledge regarding jurors’ sensitivity to various aspects of witnessing and identification conditions. Future research should establish which eyewitness factors are known to jurors and which are more problematic. Finally, it is yet unclear why expert testimony often leads to skepticism—potential explanations, such as the influence of the expert’s credentials on jurors’ judgments should be explored.

**Conclusion**

Our studies demonstrated that expert testimony did not affect mock-jurors’ verdicts. Despite improving mock-jurors’ knowledge of eyewitness factors, expert testimony did not sensitize them to suggestive identification procedures or to identification accuracy.
Furthermore, an attempt to improve the effectiveness of expert testimony by accompanying it with visual aids was not successful.

We demonstrated that mock-jurors who watched eyewitness testimony alone were more likely to believe the eyewitness than those who watched both video-recorded identification procedure and eyewitness testimony. Replicating Reardon & Fisher’s (2011) findings, mock-jurors were sensitive to identification accuracy when identification procedures were non-suggestive. This provides grounds for optimism because mock-jurors are able to differentiate between correct and false identifications. Importantly, however, this sensitivity was lost when identification procedures were suggestive, underscoring the importance of using best practice procedures. Once suggestive identification evidence appears in front of the jury—expert testimony might not be able to undo its damaging effects.
Acknowledgments

This work was partially supported by the Swinburne University Postgraduate Research Award. We would like to thank Ms Jen Owen, Mr Roy Groncki, Mr Mitchell Kloet, and Associate Professor Roger Cook.
Supplemental Material

Feedback Awareness

Mock-jurors were sensitive to the presence of feedback. When the ID procedures were suggestive, the majority (82.30%) of mock-jurors reported that the administrator commented on the eyewitness’s choice; when the ID procedures were non-suggestive, 16.90% of mock-jurors reported that the administrator made a comment after the eyewitness made an identification, $\chi^2(2) = 248.37, p < .001$. Expert testimony did not affect mock-jurors’ feedback awareness, $p > .239$.

Perceptions of the Expert

For those in expert testimony conditions, we examined whether the type of expert testimony affected mock-jurors’ perceptions of the expert. There were no significant differences between traditional expert testimony ($M = 8.25, SD = 1.32$) and expert testimony with visual aids ($M = 8.11, SD = 1.33$), $F(1, 351) = 1.04, p = .308, d = 0.11 [-0.03, 0.24]$. 
Table 13

*Percentage of Guilty Verdicts as a Function of Expert Testimony, Identification Accuracy, Type of Evidence, and Identification Procedure (Study 2)*

<table>
<thead>
<tr>
<th>Video ID + testimony</th>
<th>Control</th>
<th>Expert</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Correct</td>
<td>False</td>
<td>Total</td>
</tr>
<tr>
<td>Non-suggestive</td>
<td>51.50</td>
<td>57.10</td>
<td>54.40</td>
</tr>
<tr>
<td>Suggestive</td>
<td>64.70</td>
<td>54.10</td>
<td>57.40</td>
</tr>
<tr>
<td>Total</td>
<td>56.00</td>
<td>55.60</td>
<td>55.70</td>
</tr>
</tbody>
</table>

|                      | Correct | False  | Total   |
| Non-suggestive       | 53.70   | 42.90  | 48.70   |
| Suggestive           | 50.00   | 73.10  | 60.70   |
| Total                | 52.10   | 55.70  | 53.80   |

|                      | Correct | False  | Total   |
| Non-suggestive       | 52.70   | 50.00  | 51.40   |
| Suggestive           | 55.30   | 61.90  | 59.10   |
| Total                | 53.70   | 55.60  | 54.70   |
Table 14

Mock-jurors’ Mean Ratings for Confidence-in-Verdict as a Function of Expert Testimony, Identification Accuracy, and Identification Procedure

(Study 2)

| Video ID + testimony | Control | | Expert | | Overall | |
|---------------------|---------|-----------------|---------|-----------------|---------|
|                     | Correct | False | Total | Correct | False | Total | Correct | False | Total | Correct | False | Total |
| Video ID + testimony |         |       |       |         |       |       |         |       |       |         |       |       |
| Non-suggestive      | 5.82    | 11.74 | 8.87  | 17.29   | 10.38 | 13.7  | 10.65   | 11.16 | 10.92 | (75.04) | (76.42) | (71.34) |
|                     | (75.04) | (76.42) |       | (70.95) | (72.94) | (71.34) | (72.93) | (75.66) | (74.04) |       |       |       |
| Suggestive          | 19.47   | 6.43  | 10.54 | 30.92   | -14.35| 9.67  | 26.4    | -1.53 | 10.13 | (77.96) | (78.83) | (71.64) |
|                     | (77.96) | (78.83) |       | (67.52) | (69.85) | (71.64) | (71.14) | (76.31) | (75.13) |       |       |       |
| Total               | 10.46   | 9.01  | 9.61  | 24.38   | -1.22 | 11.71 | 17.42   | 4.87  | 10.55 | (75.53) | (77.18) | (71.15) |
|                     | (75.53) | (77.18) |       | (68.82) | (71.85) | (71.15) | (72.23) | (75.94) | (74.38) |       |       |       |
| Testimony           |         |       |       |         |       |       |         |       |       |         |       |       |
| Non-suggestive      | 8.83    | -3.77 | 3.03  | 23.04   | 13.88 | 18.64 | 14.34   | 3.41  | 9.22  | (71.82) | (71.08) | (71.82) |
|                     | (71.82) | (71.08) |       | (72.47) | (72.36) | (71.82) | (71.86) | (71.26) | (71.5)  |       |       |       |
| Suggestive          | 6.47    | 38.85 | 21.5  | 31.54   | 15.32 | 23.32 | 20.31   | 24.88 | 22.54 | (73.26) | (71.15) | (66.74) |
|                     | (73.26) | (71.15) |       | (64.06) | (69.16) | (66.74) | (68.95) | (68.28) | (68.39) |       |       |       |
| Total               | 7.83    | 14.39 | 10.86 | 28.03   | 14.76 | 21.45 | 17.33   | 14.58 | 16.01 | (71.92) | (71.43) | (68.57) |
|                     | (71.92) | (71.43) |       | (67.22) | (69.83) | (68.57) | (70.22) | (70.27) | (70.12) |       |       |       |
| Overall             |         |       |       |         |       |       |         |       |       |         |       |       |
| Non-suggestive      | 7.49    | 3.99  | 5.78  | 20.28   | 12.06 | 16.17 | 12.65   | 7.35  | 10.04 | (72.78) | (73.45) | (71.26) |
|                     | (72.78) | (73.45) |       | (71.07) | (71.94) | (71.26) | (72.08) | (73.32) | (72.59) |       |       |       |
| Suggestive          | 11.17   | 19.81 | 16.12 | 31.29   | 4.13  | 17.93 | 22.69   | 12.1  | 17.08 | (74.41) | (74.87) | (68.76) |
|                     | (74.41) | (74.87) |       | (64.97) | (70.35) | (68.76) | (69.55) | (73.19) | (71.55) |       |       |       |
| Total               | 8.92    | 11.48 | 10.26 | 26.42   | 7.7   | 17.14 | 17.37   | 9.76  | 13.49 | (73.13) | (74.1)  | (69.73) |
|                     | (73.13) | (74.1)  |       | (67.65) | (70.85) | (69.73) | (70.93) | (73.14) | (72.09) |       |       |       |
The Eyewitness Instructions and Checklist: An Alternative Method for Educating Jurors about Eyewitness Evidence

Alena Skalon\textsuperscript{a} & Jennifer L. Beaudry\textsuperscript{a}

\textsuperscript{a}Department of Psychological Sciences; School of Health Sciences; Faculty of Health, Arts and Design; Swinburne University of Technology, Hawthorn, Victoria, Australia

Corresponding author: Alena Skalon, email: alenaskalon@gmail.com. Address: Mail H24, P. O. Box 218 Hawthorn, Victoria 3122 Australia

Jennifer L. Beaudry, email: jbeaudry@swin.edu.au.

Please note that I have made small alterations to this Paper for thesis presentation. These changes were limited to placement of tables and figures, numbering of tables and figures, removing all reference lists (which are compiled at the end of the thesis), and altering some in-text citations to align with this compiled list. As no other changes were made, the referencing style and language requested by *Journal of Applied Research in Memory and Cognition* were retained.
Abstract

Traditional legal safeguards are often not effective in sensitizing mock-jurors to the quality of the eyewitness evidence. We developed the Eyewitness Instructions and Checklist (EIC) and tested how it affects mock-jurors’ evaluations of eyewitness evidence. Mock-jurors (n = 536) completed the EIC or a control activity and watched a video-recorded identification procedure and eyewitness testimony which varied in terms of the suggestiveness of the identification procedures and the identification accuracy. Mock-jurors were sensitive to identification accuracy when the identification procedures were non-suggestive, but not when they were suggestive. Although the EIC increased mock-jurors’ knowledge of eyewitness issues, it did not sensitize them to the quality of identification conditions or identification accuracy. Instead, the EIC led to overbelief. Thus, despite the sound theoretical rationale, the EIC did not prove to be an effective alternative to traditional legal safeguards. Further research should explore novel methods of educating jurors about eyewitness evidence.

Keywords: checklists, eyewitness evidence, juror decision making, video-recorded identification procedure, best practice procedures, suggestive identification procedures.
General Audience Summary

When cases involve eyewitness evidence, traditional legal safeguards (i.e., judicial instructions and expert testimony) are often not effective in sensitizing mock-jurors to the quality of the eyewitness evidence. This study tested an alternative method of educating jurors—the Eyewitness Instructions and Checklist (EIC). The EIC consists of two components: a narrated PowerPoint presentation explaining factors that can affect the reliability of eyewitness evidence, and a checklist completed at the end of the study that prompts jurors to apply information contained in the case at hand. The experimental test of the EIC demonstrated that despite increasing mock-jurors’ knowledge of eyewitness issues, it did not improve the accuracy of their assessments of eyewitness evidence. Moreover, the EIC caused mock-jurors to believe the eyewitness regardless of the actual quality of the evidence. The study also demonstrated that mock-jurors were sensitive to identification accuracy when identification procedures were non-suggestive, but not when they were suggestive. This finding highlights the importance of using best practice procedures when obtaining eyewitness identification evidence.
The Eyewitness Instructions and Checklist: An Alternative Method for Educating Jurors about Eyewitness Evidence

Criminal jury trials entrust ordinary people with life-changing decisions for others. Arguably, the criminal justice system does not effectively equip jurors with the knowledge and methods to properly evaluate the quality of the evidence, particularly in trials that involve eyewitness evidence. Psychological research repeatedly shows that traditional legal safeguards (i.e., judicial instructions and expert testimony) fail to improve jurors’ evaluations of eyewitness evidence. The current study examines the effectiveness of a new method of educating jurors—the Eyewitness Instructions and Checklist (EIC).

Jurors who must assess the quality of the evidence often lack knowledge about factors that influence the reliability of eyewitness evidence (e.g., Benton, Ross, Bradshaw, Thomas, & Bradshaw, 2006; Shaw, Garcia, & McClure, 1999). In addition to the absence of theoretical knowledge, jurors cannot differentiate between correct and mistaken identifications (Lindsay, Wells, & O’Connor, 1989; Wells, Lindsay, & Ferguson, 1979), nor can they distinguish between evidence that is likely to be reliable and unreliable (Boyce, Beaudry, & Lindsay, 2007). For example, even when mock-jurors notice that suggestive identification procedures were used, they do not integrate that information into their ultimate assessment of the evidence—the verdict (Devenport, Stinson, Cutler, & Kravitz, 2002). At the same time, there is some indication that jurors can be sensitive to eyewitness accuracy (i.e., differentiate between correct and mistaken identifications) when presented with a video-recorded identification procedure (Reardon & Fisher, 2011). However, this sensitivity might be damaged by the presence of suggestive identification procedures (Beaudry et al., 2015).

Courts around the world have recognized this issue and introduced judicial instructions or allowed expert testimony to educate the jury about factors that affect the reliability of eyewitness evidence. However, judicial instructions either do not significantly
influence mock-jurors’ judgments (e.g., Cutler, Dexter, & Penrod, 1990; Jones, Bergold, Dillon, & Penrod, 2017; Martire & Kemp, 2009) or lead to skepticism (i.e., discounting of identification evidence regardless of the quality of witnessing and identification conditions or eyewitness accuracy; e.g., Dillon, Jones, Bergold, Hui, & Penrod, 2017; Papailiou, Yokum, & Robertson, 2015). Similarly, expert testimony often results in skepticism (e.g., Cutler et al., 1990; Jones et al., 2017; Leippe, Eisenstadt, Rauch, & Seib, 2004), but sometimes leads to limited sensitivity (e.g., Cutler, Dexter, & Penrod, 1989; Devenport et al., 2002; Geiselman & Mendez, 2005).

One of the potential solutions to the ineffectiveness of traditional legal safeguards could be to improve the safeguards; however, attempts to do so were largely unsuccessful (e.g., Greene, 1988, Study 2; Ramirez, Zemba, and Geiselman, 1996, Study 2; Jones et al., 2017). For example, the Henderson judicial instructions (New Jersey v Henderson, 2011), developed after a Special Master extensively reviewed the empirical literature, provide a comprehensive overview of eyewitness factors. Thus, the Henderson instructions are a major improvement compared to most other judicial instructions that simply list some eyewitness factors (e.g., Martire & Kemp, 2009; Greene, 1988; Ramirez et al., 1996). Jones et al. (2017) further enhanced the Henderson instructions and provided an explanation for each of the eyewitness factors. Despite these improvements, neither version of the Henderson instructions influenced mock-jurors’ judgments (Jones et al., 2017). Their results suggest that even the modified judicial instructions are not an effective method of educating the jury.

Efforts to improve expert testimony have produced mixed results. Geiselman et al. (2002, Experiment 1) presented mock-jurors with specific expert testimony that discussed the presence of various eyewitness factors in the case at hand (e.g., the expert explained that cross-race identifications are 10-15% less likely to be accurate and noted that the eyewitness and perpetrator were from different races). Compared to general expert testimony that
induced skepticism, the specific testimony sensitized mock-jurors to witnessing and identification conditions (i.e., mock-jurors believed the witness more often when witnessing and identification conditions were good compared to poor). However, in another study, specific expert testimony lead to the same outcome—skepticism—as general expert testimony (Fox & Walters, 1986). Taken together, expert testimony might be more promising as a legal safeguard in eyewitness cases than judicial instructions, especially if it is tailored to a particular case. However, it is very unlikely that an eyewitness expert will be allowed to ‘enter the province of the jury’ by evaluating the eyewitness evidence and commenting on the probable reliability of the identification (Woller, 2003; Mnookin, 2015). Thus, researchers should develop a method to present information similar to that provided by the expert and equip the jury with a tool to apply that information to the case at hand.

With that goal in mind, Wise, Fishman, and Safer (2009) created the Interview-Identification-Eyewitness factors (I-I-Eye) educational aid. The I-I-Eye consists of three steps: assessment of the interview, assessment of the identification procedure, and assessment of witnessing conditions. Each assessment consists of several questions presented in a PowerPoint presentation (24 slides in total) prior to eyewitness evidence. Pawlenko, Safer, Wise, and Holfeld (2013) tested the effectiveness of the I-I-Eye by comparing it with a Neil v Biggers educational aid (1972; describing five relevant eyewitness factors) and a control condition that did not present any information about eyewitness evidence. Only participants who watched the I-I-Eye were sensitive to identification conditions: they were more likely to render a guilty verdict when the identification conditions were non-suggestive than suggestive. Safer et al. (2016) conducted a similar study, but with the addition of circumstantial evidence. Again, the I-I-Eye sensitized participants to the quality of identifications conditions. The success of the I-I-Eye suggests that alternative methods of educating jurors can be more effective than traditional legal safeguards. Given that they used
simulated eyewitness evidence, however, it is unclear as to whether the I-I-Eye would sensitize jurors to identification accuracy.

Building on the potential effectiveness of the I-I-Eye, we created a new educational tool for use in cases with eyewitness evidence—the Eyewitness Instructions and Checklist (EIC)—that consists of two components. The first component is a narrated PowerPoint presentation that jurors watch before the case. Rather than reading the slides like in the I-I-Eye, jurors watch the presentation that includes graphs and pictures illustrating the key points of the material. We included visual information into the EIC because expert testimony research in areas other than eyewitness psychology demonstrated that supplementing expert testimony with visual aids can improve mock-jurors’ understanding and application of the information (Jones & Kovera, 2015; Park & Feigenson, 2013).

The second component of the EIC is a checklist that jurors complete after reviewing the case material. This is also different from the I-I-Eye because the EIC presents information both before and after the evidence. The EIC checklist was inspired by question trails that are already used around the world (Clough, 2013), as well as by promising results showing that a flowchart/question trail better educated jurors about the validity of scientific studies (Jones & Kovera, 2015). The EIC checklist, however, cannot be categorized as either a flowchart or a question trail because it does not guide jurors toward accepting or rejecting the evidence. Rather, it acts as a recall aid and asks jurors to consider which eyewitness factors were present in the case and how each might have affected the eyewitness (see materials section for a more detailed description of the checklist). Thus, the checklist might help mock-jurors to integrate and apply information that was provided in the presentation. If effective, the EIC could reduce the need for judicial instructions or expert testimony and could be used in modules to address issues relevant to a specific case (Mnookin, 2015).
Present Study

The present study tested the effectiveness of the EIC in sensitizing jurors to suggestive identification procedures and to identification accuracy. Unlike the majority of studies that used simulated testimony in their tests of legal safeguards and educational aids, we used a genuine eyewitness paradigm (Wells et al., 1979). Mock-jurors watched video-recorded testimony and identification evidence from an eyewitness who saw a mock-crime and attempted to identify the perpetrator. In addition, this study compared the performance of students and community members and by doing so addressed one of the common criticisms articulated by legal scholars—the use of students in jury simulation research (e.g., Leverick, 2016).

We preregistered five key hypotheses for this study (https://tinyurl.com/y7u829jm). We predicted a main effect of identification accuracy (hypothesis 1): mock-jurors who watch correct identifications will be more likely to render guilty verdicts and to believe that the witness correctly identified the culprit than mock-jurors who watch mistaken identifications. We also hypothesized that this main effect will be qualified by an Identification Accuracy × Identification Procedure interaction (hypothesis 2). When viewing a non-suggestive procedure, mock-jurors will be sensitive to identification accuracy, rendering more guilty verdicts and estimating a higher likelihood that the witness correctly identified the culprit when they view a correct identification (cf. mistaken). When viewing a suggestive procedure, mock-jurors’ belief of the eyewitnesses will be similar regardless of identification accuracy.

We expected a significant Educational Aid × Identification Procedure interaction on mock-jurors’ perceptions of the quality of identification conditions (hypothesis 3). When they view an identification obtained from a suggestive procedure, mock-jurors who completed the EIC (cf. control) will be significantly less likely to render a guilty verdict, less likely to believe that the witness correctly identified the culprit, and perceive the administrator as
having more influence. When they view a non-suggestive procedure, two potential outcomes are possible. We predict that it is most likely that there will be no significant differences between mock-jurors who completed the EIC and those in the control condition. However, if the EIC is an effective safeguard, the EIC should improve sensitivity to non-suggestive identification conditions. If that is the case, mock-jurors who watched a non-suggestive identification and completed the EIC will be more likely to render a guilty verdict, more likely to believe that the witness correctly identified the culprit, and assign lower ratings of administrator influence than mock-jurors in the control condition.

Two studies have demonstrated that expert testimony did not improve mock-jurors’ sensitivity to the accuracy of genuine eyewitnesses (Wells, Lindsay, & Tousignant, 1980; Martire & Kemp, 2009), but whether other educational aids can sensitize jurors to eyewitness accuracy is still unknown. If the EIC improves mock-jurors’ ability to differentiate between correct and mistaken identifications, a significant Educational Aid × Identification Accuracy interaction should emerge (hypothesis 4). In the control condition, mock-jurors who watch a correct identification will be significantly more likely to render a guilty verdict and be more likely to believe that the witness correctly identified the culprit compared to those who watched a mistaken identification (e.g., Reardon & Fisher, 2011). For mock-jurors who completed the EIC, we expect that this pattern (i.e., greater belief for correct identifications than for mistaken ones) will be exaggerated. We predict that this exaggerated sensitivity will mostly be driven by the difference in their belief of mistaken eyewitnesses. That is, when they view a mistaken identification, mock-jurors in the control condition will be significantly more likely to believe the eyewitness compared to those who completed the EIC. Finally, we predict that mock-jurors in the EIC condition will demonstrate better knowledge of eyewitness factors compared to mock-jurors in the control condition (hypothesis 5).
Method

Disclosures

The data for the study (both raw and clean) are available on the Open Science Framework (OSF; https://tinyurl.com/y7u829jm). We report all data exclusions and manipulations in the study. Please see our OSF project for all measures and the details regarding how we determined our sample size. This project has been approved by Swinburne’s Human Research Ethics Committee (SUHREC) in line with the Australian National Statement on Ethical Conduct in Human Research.

Design

We used a 2 (Identification [ID] Procedure: suggestive vs. non-suggestive) × 2 (ID Accuracy: correct vs. mistaken) × 2 (Educational Aid: EIC vs. control) between-subjects factorial design.

Participants

The final sample consisted of 694 mock-jurors (536 students and 158 community members). The student participants (132 males, 402 females, 2 other) were enrolled in first-year on-campus and online psychology and criminology units at a large Australian university and participated in exchange for extra credit. The community participants were members of Prolific Academic (77 males, 80 females, 1 other; Prolific, Oxford, UK). Participants’ age ranged from 16 to 74 years and did not significantly differ across the samples (students: $M = 30.32, SD = 10.58$; community members: $M = 30.89, SD = 10.40$), $t(691) = -0.60, p = .549$. Students self-identified as Caucasian (78%), followed by other (11%), Asian (9%), and Indigenous (2%). Similarly, community members self-identified as Caucasian (77%), followed by other (17%), Asian (5%), and Indigenous (1%).
Materials

Eyewitness Instructions and Checklist. The EIC consists of two components: a video (narrated PowerPoint presentation; duration of the video is 14 min 32 s) and a checklist. The video presentation first provided an overview of the three memory stages and the research methods used by eyewitness researchers; in addition, the video explained the principles of identification procedures and the differences between a suspect, a perpetrator, and a filler. The focus of the presentation, however, was on eyewitness factors separated into three categories: (i) factors at the time of the crime: duration, distance, lighting, presence of a weapon, intoxication, disguises, stress, and cross-race identification; (ii) factors between the criminal event and the identification procedure: delay and changed appearance; and (iii) factors at the time of the identification: lineup administration (single-blind or double-blind), lineup instructions, post-identification feedback, multiple viewings of the suspect, and the relationship between confidence and accuracy.

After the video, mock-jurors saw a completed sample checklist based on the facts of a fictional case. Mock-jurors were informed that the completed checklist was for training purposes only. The second part—the checklist for the case at hand—was presented at the end of the study (see our OSF project for the full script of the EIC video and the checklist). The checklist asked mock-jurors to consider whether each of the factors was present in the case they watched and they could tick: “yes,” “no,” or “unknown.” After completing the checklist, mock-jurors received the following instructions:

Please remember that it is up to you to decide whether and to what extent these factors affected the witness and to evaluate the identification evidence in combination with all of the evidence in this case. Keeping in mind the presence or absence of the impairing factors listed above, how would you evaluate the overall quality of eyewitness identification evidence in this case? Please explain in your own words.
Control condition. Given that the EIC consists of two components, we separated activities in the control condition into two parts. At the beginning of the study, mock-jurors in the control condition watched a jury duty video of comparable duration to the EIC (12 min 36 s) and completed a brief questionnaire that tested their understanding of the information presented in the video (the questionnaire acted as a control activity and was not assessed). The second component was presented at the end of the study and acted as a control for the checklist. Mock-jurors were asked to describe what factors they thought were important in the case and the strengths and weaknesses of the case.

Eyewitness identification and eyewitness testimony stimuli. We aimed to recreate the circumstances of a trial by showing mock-jurors videos of genuine eyewitnesses. Because the stimuli were produced in a laboratory, we know the ground truth of whether the witness was correct or mistaken; thus, we can examine the actual accuracy of mock-jurors’ belief decisions. In the genuine witness paradigm, eyewitness-participants—as opposed to eyewitness-actors who read a script—first witness a crime, attempt to identify the culprit, and testify about their experience (Wells et al., 1979). We used a subset of stimuli produced as part of a larger project (described in Beaudry et al., 2015). To increase stimulus sampling (Wells & Windschitl, 1999), we used a total of 16 eyewitness videos in this study with four videos per each identification condition (i.e., four non-suggestive mistaken identifications, four non-suggestive correct identifications, four suggestive mistaken identifications, four suggestive correct identifications); however, each mock-juror participant watched only one eyewitness. We collapsed data across eyewitnesses in each of the conditions (e.g., Beaudry et al., 2015; Charman & Wells, 2007).

Eyewitnesses made their identification under non-suggestive or suggestive conditions. In the non-suggestive identification procedure condition, the lineup administrator did not know the identity of the suspect, the administrator provided unbiased instructions warning
that the culprit may or may not be present in the lineup, and the lineup was constructed to be fair with no lineup member standing out from the others (used by Beaudry et al., 2015). In the suggestive identification procedure condition, the administrator knew the identity of the suspect, provided instructions saying, “All you have to do is pick him out,” and the lineup was constructed to be biased such that the suspect clearly stood out from other lineup members. Correct identifications were obtained from eyewitnesses who viewed a culprit-present lineup and correctly identified the culprit, whereas mistaken identifications were obtained from eyewitnesses who viewed a culprit-absent lineup and identified an innocent suspect.

**Questionnaire.** Please refer to our OSF project for the full questionnaire. Our main dependent variables were verdict (binary: guilty vs. not guilty); the likelihood that the person identified by the witness was the culprit [ID likelihood], ranging from 0% (*very unlikely*) to 100% (*very likely*); lineup fairness, ranging from 0 (*completely unfair*) to 10 (*completely fair*); and two aggregate measures. The perceptions of the witness measure was created by averaging mock-jurors’ ratings of the witness’s trustworthiness, reliability, confidence, persuasiveness, and their perceptions of the witness’s view of the culprit, the amount of attention the witness paid to the culprit’s face, and the clarity of the witness’s memory at the time of the lineup, $rs > .39, ps < .001$. This scale ranged from 0 (*not reliable*) to 10 (*very reliable*). The second combined measure—perceptions of the administrator—was created by averaging four questions assessing mock-jurors’ perceptions of the administrator’s influence on the eyewitness’s confidence, knowledge of the suspect’s identity, pressure to choose someone at the time of the lineup, and the administrator’s investment in the outcome of the trial $rs > .42, ps < .001$. This combined measure ranged from 0 (*no influence*) to 10 (*definite influence*).
Knowledge of eyewitness factors. Eleven questions assessed mock-jurors’ knowledge of eyewitness factors; five questions had true/false response options and six questions had multiple-choice response options (with one correct option).

Manipulation checks. Manipulation checks asked if mock-jurors watched the videos. In response to concerns that many Mechanical Turk “participants” were actually bots (Dreyfuss, 2018; Stokel-Walker, 2018), we introduced additional, simple attention checks in August 2018 (e.g., 2 + 2 = ?). We were more concerned about the validity of our community participants obtained from Prolific Academic than our student population obtained from our department’s Research Experience Program, so we introduced two attention checks for students and four attention checks for community members. Given that we introduced these attention checks after we had already started data collection, only a portion of our participants completed them (30% of student participants; 56% of community members).

Procedure

Participants who signed up to participate received a link to the online Qualtrics study (Qualtrics, Provo, UT), and were directed to a page with an information statement. Those who agreed to participate were informed that they will be presented with trial materials. Mock-jurors in the experimental condition watched a narrated PowerPoint video presentation (the first component of the EIC) and saw an example of a completed sample checklist. Mock-jurors in the control condition watched a video about jury duty and completed a series of filler questions. Following this, they read a case summary of a fictional theft case, read opening statements of two attorneys, watched a video of an eyewitness making an identification from a lineup, and watched the testimony of the same eyewitness (as mentioned, each participant watched the identification attempt and testimony of one of the 16 eyewitnesses). After watching the eyewitness videos, mock-jurors read the attorneys’ closing statements and either completed the checklist (the second component of the EIC) or
completed the control activity (i.e., described the strengths and weaknesses of the case). At the end of the study, all mock-jurors completed the questionnaire evaluating their perceptions of the eyewitness evidence, completed the manipulation and attention checks, and viewed a debriefing statement.

**Results**

**Data Preparation**

**Manipulation checks.** The initial sample consisted of 965 participants (776 students and 189 community members). We excluded data from 113 participants who did not complete the study. We excluded participants who completed the study too quickly ($n = 25$) or those who took too long to complete the study ($n = 98$). We determined the minimum durations for the control condition (less than 29 min) and the EIC condition (less than 31 min) based on the combined duration of all videos in that condition. We determined the maximum duration for the control (90 min) and EIC conditions (112 min) based on extremes identified in the stem-and-leaf diagram. Additional participants were excluded because they indicated that they did not see one of the videos ($n = 12$) or they failed at least one of the attention checks ($n = 23$). The final sample consisted of 694 participants.

**Confidence-in-verdict measure.** To create a confidence-in-verdict measure we multiplied each mock-juror’s confidence rating by -1 for a not-guilty verdict and by +1 for a guilty verdict (Tenney, MacCoun, Spellman, & Hastie, 2007). Thus, a score of -100% indicates that a participant was very confident in a not-guilty verdict, whereas a score of 100% indicates high confidence in a guilty verdict.

**Data Analytic Approach**

We analyzed the binary verdict measure with a binary logistic regression, with the main effects and all relevant interactions entered hierarchically. We report odds ratios and percentages for this measure. We used ANOVAs to examine the independent and interactive
effects of ID accuracy, ID procedure, and educational aid on the confidence-in-verdict, the likelihood that the person identified by the witness was the culprit, perceptions of the witness, perceptions of lineup fairness, and perceptions of the lineup administrator. For these continuous measures, we report means and standard deviations. We report 95% confidence intervals for our effect sizes in square brackets. We used chi-square analyses to examine the effect of the educational aid on mock-jurors’ knowledge of eyewitness factors (for the full questionnaire see our OSF page). We also examined continuous variables using Bayesian analysis to investigate whether it supports the findings obtained with the null-hypothesis significance testing (JASP Team, 2018); these results are available in supplementary materials.

Are there Differences between Students and Community Members?

Given the recent meta-analytic results of Bornstein et al. (2017), we did not expect significant differences between students and community members. Nonetheless, we wanted to ensure that there were no differences between our two samples in terms of their verdicts or ID likelihood. We preregistered that we would look for differences between our two samples, but we did not have any hypotheses about how sample type would interact with other independent variables.

We ran a binary logistic regression to test whether our samples differed in their verdict decisions (see Table 18). The binary logistic regression revealed that the model that included all main effects and interactions between the sample type and ID accuracy, ID procedure, the educational aid was statistically significant, $\chi^2(7) = 49.98, p < .001$. However, sample type was not significant on its own or as an interaction term, $p_s > .073$. We also ran an ANOVA to test whether there are differences between students and community members in their responses to the ID likelihood measure. Two interactions were statistically significant. A significant Sample $\times$ Educational Aid interaction, $F(1, 678) = 4.34, p = .038, \eta^2_p = 0.01,$
demonstrated that students who watched suggestive ID procedures ($M = 54.73$, $SD = 23.77$) were more likely to believe that the witness correctly identified the culprit compared to students who watched non-suggestive IDs ($M = 50.85$, $SD = 21.05$), $t(534) = 2.00$, $p = .046$, $d = 0.17 [-1.72, 2.07]$. On the other hand, community members believed the witness regardless of the suggestiveness of ID procedures (suggestive: $M = 54.34$, $SD = 22.07$; non-suggestive: $M = 52.57$, $SD = 24.69$), $t(156) = 0.48$, $p = .635$, $d = 0.08 [-3.55, 3.70]$. A significant Sample $\times$ ID Procedure $\times$ Educational Aid interaction has also emerged, $F(1, 678) = 5.43$, $p = .020$, $\eta^2_p = 0.01$. However, follow-up tests were not statistically significant.

Due to the differences found for our key measure—ID likelihood—we need to analyze the two samples separately. Given that we did not expect differences and did not intend to compare our samples, we did not collect equal numbers of students ($n = 536$) and community members ($n = 158$). Importantly, the sample size for community members was too small to examine in isolation for this type of design. Thus, for the remainder of the manuscript, we report the results obtained only from the student sample because this is the larger sample. We report the results for the community members in the supplementary materials, but urge that they are interpreted with caution given the smaller sample size.

**Hypothesis 1: Are Mock-jurors Sensitive to Eyewitness Accuracy?**

In line with our hypothesis, mock-jurors who watched correct IDs were more likely to render a guilty verdict (42%) than those who watched mistaken IDs (24%), $B = 1.59$, $SE = .30$, Wald ($1, n = 536$) = 27.10, $p < .001$, OR = 4.88 [2.70, 8.80] (see Table 19). Furthermore, mock-jurors who watched correct IDs were less confident in a not guilty verdict, $d = 0.48 [-4.88, 5.85]$, were more likely to believe that the witness correctly identified the culprit, $d = 0.50 [-1.35, 2.34]$, perceived the witness to be more reliable, $d = 0.29 [-0.19, 0.77]$, and perceived the administrator as having less influence on the witness, $d = 0.20 [0.01, 0.39]$, compared to mock-jurors who watched mistaken IDs. The only measure that was not affected
by ID accuracy was lineup fairness—mock-jurors tended to rate the lineup as fair regardless of ID accuracy (see Table 15 for ANOVA significance tests and Table 16 for descriptive statistics). The large standard deviations for the confidence-in-verdict measure are caused by the large variability of the scale (i.e., from -100 to 100) and are typical (e.g., Palmer, Button, Barnett, & Brewer, 2014; Tenney et al., 2007).
Table 15

Univariate Analyses of Variance Results for ID Procedure × ID Accuracy × Educational Aid Analyses on Continuous Dependent Variables (N = 536)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Confidence-in-verdict</th>
<th>ID likelihood</th>
<th>Lineup fairness</th>
<th>Perceptions of administrator</th>
<th>Perceptions of eyewitness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$F$</td>
<td>$\eta^2$</td>
<td>$F$</td>
<td>$\eta^2$</td>
<td>$F$</td>
</tr>
<tr>
<td>ID accuracy</td>
<td>34.80***</td>
<td>0.06</td>
<td>35.25***</td>
<td>0.06</td>
<td>1.11</td>
</tr>
<tr>
<td>ID procedure</td>
<td>4.63*</td>
<td>0.01</td>
<td>5.01*</td>
<td>0.01</td>
<td>103.43***</td>
</tr>
<tr>
<td>Educational aid</td>
<td>8.04**</td>
<td>0.02</td>
<td>5.28*</td>
<td>0.01</td>
<td>2.05</td>
</tr>
<tr>
<td>ID Procedure × ID Accuracy</td>
<td>9.13**</td>
<td>0.02</td>
<td>12.49***</td>
<td>0.02</td>
<td>0.20</td>
</tr>
<tr>
<td>ID Procedure × Educational Aid</td>
<td>1.03</td>
<td>0.00</td>
<td>0.49</td>
<td>0.00</td>
<td>6.10*</td>
</tr>
<tr>
<td>ID Accuracy × Educational Aid</td>
<td>2.07</td>
<td>0.00</td>
<td>0.36</td>
<td>0.00</td>
<td>0.08</td>
</tr>
<tr>
<td>ID Procedure × ID Accuracy × Educational Aid</td>
<td>1.07</td>
<td>0.00</td>
<td>0.01</td>
<td>0.00</td>
<td>0.53</td>
</tr>
</tbody>
</table>

*Note. All dfs = 1.*

***p < .001. **p < .01. *p < .05.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Control</th>
<th></th>
<th>EIC</th>
<th></th>
<th>Overall</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Correct ID</td>
<td>Mistaken ID</td>
<td>Total</td>
<td>Correct ID</td>
<td>Mistaken ID</td>
<td>Total</td>
</tr>
<tr>
<td>Confidence-in-verdict</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-suggestive</td>
<td>-12.12 (67.91)</td>
<td>-47.10 (50.35)</td>
<td>-29.96 (61.94)</td>
<td>11.17 (65.79)</td>
<td>-50.63 (40.77)</td>
<td>-18.00 (63.29)</td>
</tr>
<tr>
<td>Suggestive</td>
<td>-16.74 (67.43)</td>
<td>-30.14 (64.56)</td>
<td>-22.81 (66.24)</td>
<td>6.35 (69.41)</td>
<td>-11.46 (65.97)</td>
<td>-3.81 (67.76)</td>
</tr>
<tr>
<td>Total</td>
<td>-14.46 (67.48)</td>
<td>-39.47 (57.59)</td>
<td>-26.53 (64.03)</td>
<td>9.04 (67.16)</td>
<td>-29.52 (58.93)</td>
<td>-11.02 (65.78)</td>
</tr>
<tr>
<td>ID likelihood</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-suggestive</td>
<td>57.66 (21.77)</td>
<td>41.04 (18.72)</td>
<td>49.19 (21.86)</td>
<td>61.71 (18.82)</td>
<td>42.95 (16.24)</td>
<td>52.86 (19.93)</td>
</tr>
<tr>
<td>Suggestive</td>
<td>53.88 (24.53)</td>
<td>50.56 (24.95)</td>
<td>52.37 (24.69)</td>
<td>60.65 (23.88)</td>
<td>55.00 (21.21)</td>
<td>57.43 (22.47)</td>
</tr>
<tr>
<td>Total</td>
<td>55.75 (23.21)</td>
<td>45.32 (22.18)</td>
<td>50.71 (23.27)</td>
<td>61.24 (21.11)</td>
<td>49.45 (19.94)</td>
<td>55.11 (21.30)</td>
</tr>
<tr>
<td>Lineup fairness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-suggestive</td>
<td>7.86 (2.37)</td>
<td>7.96 (2.28)</td>
<td>7.91 (2.32)</td>
<td>8.34 (2.06)</td>
<td>7.97 (2.13)</td>
<td>8.17 (2.09)</td>
</tr>
<tr>
<td>Suggestive</td>
<td>6.33 (3.04)</td>
<td>5.87 (3.22)</td>
<td>6.12 (3.12)</td>
<td>5.30 (3.21)</td>
<td>5.06 (3.09)</td>
<td>5.17 (3.13)</td>
</tr>
<tr>
<td>Total</td>
<td>7.09 (2.83)</td>
<td>7.02 (2.92)</td>
<td>7.06 (2.87)</td>
<td>7.01 (3.03)</td>
<td>6.40 (3.05)</td>
<td>6.69 (3.05)</td>
</tr>
<tr>
<td>Perceptions of the adminstrator</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-suggestive</td>
<td>2.11 (1.88)</td>
<td>2.90 (1.95)</td>
<td>2.51 (1.95)</td>
<td>2.22 (1.77)</td>
<td>2.52 (1.74)</td>
<td>2.36 (1.75)</td>
</tr>
<tr>
<td>Suggestive</td>
<td>3.97 (2.27)</td>
<td>4.15 (2.50)</td>
<td>4.05 (2.37)</td>
<td>3.85 (2.56)</td>
<td>4.19 (2.27)</td>
<td>4.04 (2.39)</td>
</tr>
<tr>
<td>Total</td>
<td>3.04 (2.28)</td>
<td>3.46 (2.29)</td>
<td>3.24 (2.29)</td>
<td>2.95 (2.30)</td>
<td>3.42 (2.20)</td>
<td>3.20 (2.25)</td>
</tr>
<tr>
<td>Perceptions of the eyewitness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-suggestive</td>
<td>6.35 (2.10)</td>
<td>4.05 (1.63)</td>
<td>5.18 (2.20)</td>
<td>6.72 (1.73)</td>
<td>4.45 (1.17)</td>
<td>5.65 (1.87)</td>
</tr>
<tr>
<td>Suggestive</td>
<td>6.15 (2.01)</td>
<td>5.23 (1.99)</td>
<td>5.73 (2.04)</td>
<td>7.16 (1.53)</td>
<td>5.94 (1.62)</td>
<td>6.46 (1.69)</td>
</tr>
<tr>
<td>Total</td>
<td>6.25 (2.05)</td>
<td>4.58 (1.89)</td>
<td>5.44 (2.14)</td>
<td>6.91 (1.65)</td>
<td>5.25 (1.60)</td>
<td>6.05 (1.83)</td>
</tr>
</tbody>
</table>

Note. Standard deviations are reported in parentheses.
Hypothesis 2: Does the Suggestiveness of Identification Procedures Affect Mock-jurors’ Sensitivity to Identification Accuracy?

Importantly, the main effect of ID accuracy should be interpreted within its significant interaction with ID procedure for verdict and continuous measures. Consistent with our hypothesis, the ID Accuracy × ID Procedure interaction on verdict was significant, $B = -1.29$, $SE = .40$, Wald $(1, n = 536) = 10.50$, $p = .001$, OR = 0.28 [0.13, 0.60]. Of the mock-jurors who watched non-suggestive IDs, 44% rendered guilty verdicts for correct IDs whereas 14% rendered guilty verdicts for mistaken IDs, $\chi^2(1) = 30.58$, $p < .001$, OR = 4.90 [2.72, 8.82]. Of the mock-jurors who watched suggestive IDs, 39% rendered guilty verdicts for correct IDs and 33% rendered guilty verdicts for mistaken IDs, $\chi^2(1) = 0.92$, $p = .336$, OR = 1.28 [0.77, 2.13].

We found significant results for three of our continuous measures. In the non-suggestive condition, mock-jurors who watched correct IDs were less confident in their not guilty verdicts, $t(274) = 6.78$, $p < .001$, $d = 0.82$ [-6.02, 7.66], and more likely to believe that the witness correctly identified the culprit, $t(274) = -7.69$, $p < .001$, $d = 0.93$ [-1.32, 3.18], compared to mock-jurors who watched mistaken IDs. This sensitivity was lost when ID procedures were suggestive, $ts < 1.56$, $ps > .120$. Similarly, when IDs were non-suggestive, mock-jurors rated the eyewitness as more reliable when IDs were correct compared to mistaken, $t(274) = -11.13$, $p < .001$, $d = 1.35$ [1.15, 1.55]. Unexpectedly, though, this significant effect also emerged when IDs were suggestive, $t(258) = -4.16$, $p < .001$, $d = 0.52$ [0.29, 0.74]. The ID Procedure × ID Accuracy did not affect mock-jurors’ perceptions of lineup fairness or perceptions of the administrator.

Hypotheses 3 and 4: Does the EIC Sensitize Mock-jurors to Identification Conditions or to Identification Accuracy?
We predicted that mock-jurors who completed the EIC will be more sensitive to ID conditions and/or more sensitive to identification accuracy than those in the control condition. Educational aid had a significant effect on verdict, $B = 0.37$, $SE = .19$, Wald $(1, n = 536) = 3.81, p = .051$, OR = 1.45 [1.00, 2.12], but contrary to our expectations, the EIC led to overbelief. A larger percentage of mock-jurors voted guilty if they completed the EIC (37%) than the control task (29%). Mock-jurors in the control condition were also more confident in their not-guilty verdicts than those in the EIC condition, $d = 0.24 [-.50, .51]$. Finally, mock-jurors in the EIC condition estimated a greater likelihood that the witness correctly identified the culprit, $d = 0.20 [-.17, 2.09]$, and perceived the witness to be more reliable, $d = 0.31 [.14, .47]$, than those in the control condition.

The only indication of improved sensitivity as a result of completing the EIC was evident in mock-jurors’ perceptions of lineup fairness. A significant Educational Aid $\times$ ID Procedure interaction indicated that mock-jurors in both the control, $t(288) = 5.58, p < .001$, and the EIC condition, $t(244) = 8.87, p < .001$, were sensitive to the suggestiveness of ID procedures. Mock-jurors in the control condition who watched an ID obtained from a non-suggestive procedure estimated higher ratings of lineup fairness than those who watched a suggestive ID, $d = 0.66 [0.34, 0.97]$. This sensitivity to suggestiveness was exaggerated for mock-jurors who watched the EIC, $d = 1.14 [0.81, 1.47]$. The educational aid manipulation had no effect on mock-jurors’ perceptions of the lineup administrator.

**Hypothesis 5: Does the EIC Improve Mock-jurors’ Knowledge of Eyewitness Factors?**

We expected that mock-jurors who completed the EIC would demonstrate better knowledge of eyewitness factors than mock-jurors in the control condition. Our hypothesis was supported: mock-jurors who completed the EIC were more knowledgeable about eyewitness factors (cf. control) on 8 out of 11 questions (see Table 17). These results indicate that mock-jurors paid attention to and understood the information provided in the EIC.
Table 17

Percentage of Correct Answers to Knowledge Questions in Control and the EIC Conditions (Student Sample)

<table>
<thead>
<tr>
<th>Question</th>
<th>Correct answer</th>
<th>Control</th>
<th>EIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. What are the stages of memory that are involved in the process of remembering?</td>
<td>Encoding; Retention; Retrieval</td>
<td>75%</td>
<td>98%***</td>
</tr>
<tr>
<td>2. A witness may think that s/he is accurate and still be mistaken.</td>
<td>True</td>
<td>99%</td>
<td>99%</td>
</tr>
<tr>
<td>3. Stress increases the accuracy of a witness’ identification.</td>
<td>False</td>
<td>59%</td>
<td>80%***</td>
</tr>
<tr>
<td>4. Which of these factors should not be considered when evaluating the time of the crime?</td>
<td>Gender of an eyewitness</td>
<td>85%</td>
<td>94%***</td>
</tr>
<tr>
<td>5. Presence of a weapon:</td>
<td>Decreases the likelihood that the witness will make an accurate identification</td>
<td>63%</td>
<td>74%**</td>
</tr>
<tr>
<td>6. Choose the most accurate statement:</td>
<td>Eyewitness’s memories can change because of discussions with other people</td>
<td>87%</td>
<td>89%</td>
</tr>
<tr>
<td>7. Which of these factors should not be considered when evaluating the identification procedure?</td>
<td>Whether the eyewitness was tired when the photoboard was conducted</td>
<td>54%</td>
<td>82%***</td>
</tr>
<tr>
<td>8. A photoboard in which the suspect stands out is more reliable because it is easier to choose the suspect</td>
<td>False</td>
<td>72%</td>
<td>89%***</td>
</tr>
<tr>
<td>9. Post-identification feedback improve the reliability of eyewitness testimony</td>
<td>False</td>
<td>68%</td>
<td>80%**</td>
</tr>
<tr>
<td>10. Post-identification feedback affects:</td>
<td>Confidence in the identification decision</td>
<td>76%</td>
<td>79%</td>
</tr>
<tr>
<td>11. Photoboards should be banned because they are utterly suggestive.</td>
<td>False</td>
<td>73%</td>
<td>94%***</td>
</tr>
</tbody>
</table>

*p < .05. **p < .01. ***p < .001.
Discussion

For this study, we recruited students and community members. Unlike Bornstein et al.’s (2017) meta-analysis, but similarly to Jones et al.’s (2017) findings, significant differences between the two samples emerged; thus, we analyzed them separately and focused this paper on the larger sample of student participants. Our discussion will similarly be limited to the results obtained from our student sample. Our first hypothesis—that mock-jurors would be sensitive to ID accuracy—was strongly supported. When IDs were correct, compared to mistaken, mock-jurors were more likely to render a guilty verdict, more likely to believe that the witness correctly identified the culprit, and perceived the witness to be more reliable.

Although this finding contrasts previous research demonstrating mock-jurors’ insensitivity to eyewitness accuracy (Lindsay et al., 1989; Wells et al., 1979), it is important to note that our mock-jurors saw both the eyewitness’s testimony and a video-record of their identification procedure. Thus, our results lend further support for the argument that providing mock-jurors with video-recorded identification testimony plus eyewitness testimony allows mock-jurors to differentiate between correct and mistaken eyewitnesses (Reardon & Fisher, 2011).

Importantly, the main effect of ID accuracy should be interpreted within its significant interaction with ID procedure. As predicted (hypothesis 2), mock-jurors demonstrated sensitivity to identification accuracy—as measured by verdicts and estimates that the eyewitness correctly identified the culprit—only when identifications were obtained from non-suggestive identification procedures. Mock-jurors were not sensitive to ID accuracy when ID procedures were suggestive. This finding is in line with Beaudry et al. (2015) and with Smalarz and Wells (2014). This effect emerged despite mock-jurors rating the lineup as less fair and the administrator as having more influence on the eyewitness when identification procedures were suggestive (cf. non-suggestive; e.g., Devenport, et al., 2002).
We also hypothesized that the EIC would sensitize mock-jurors’ to the quality of identification procedures (hypothesis 3) and to eyewitness accuracy (hypothesis 4). Neither of these hypotheses were supported. Moreover, the EIC led to overbelief. Mock-jurors who watched the EIC were more likely to render guilty verdicts, were less confident in not guilty verdicts, estimated greater likelihood that the witness correctly identified the culprit, and perceived the witness as more reliable than mock-jurors in the control condition. This is a very unusual finding. The only study, to our knowledge, which found that a legal safeguard (i.e., the Telfaire instructions) led to overbelief presented the judicial instructions both before and after the eyewitness evidence (Ramirez et al., 1996). The EIC is also presented before (i.e., PowerPoint presentations) and after (i.e., the checklist) the evidence. The timing might explain this result.

Another potential explanation is that the EIC contained information about multiple eyewitness factors (e.g., lighting, presence of a weapon, stress, and changed appearance). In our mock-case, the problematic factors were limited to the suggestiveness of the identification procedure, and all estimator variables were relatively good (e.g., short distance, good lighting, low stress). Thus, one possibility is that when mock-jurors completed the checklist and evaluated the strength of the eyewitness evidence against each of the factors, they found that the evidence had very few issues, without potentially realizing the importance and power of these few procedural issues. Indirectly supporting this explanation is the fact that mock-jurors in the EIC condition rated the suggestive lineup as less fair than the non-suggestive lineup. It is worth noting, however, that despite the fact that the EIC technically led to overbelief, conviction rates were still low—only 44% of the mock-jurors rendered a guilty verdict even when identifications were correct and were obtained using non-suggestive procedures. Our case might have been too weak to demonstrate real differences between the control and the EIC condition.
Finally, in line with our prediction, the EIC improved mock-jurors’ knowledge of eyewitness factors (hypothesis 5). Mock-jurors who watched the EIC demonstrated a significantly better understanding of 8 out of 11 eyewitness factors. Regardless of the presence of the EIC, mock-jurors demonstrated a perfect understanding that a witness may think that s/he is accurate and still be mistaken, indicating that this question was too general in nature. Furthermore, regardless of the educational aid condition, mock-jurors demonstrated a good understanding of the circumstances under which eyewitnesses’ memories can change and the influence of post-identification feedback on eyewitnesses’ confidence.

Overall, the EIC did not meet expectations—it did not improve mock-jurors’ sensitivity to identification procedures or to identification accuracy. Instead, in contrast to our predictions and to previous research, it led to overbelief. This is despite the fact that mock-jurors clearly paid attention and understood the EIC—which is evident in improved knowledge of eyewitness factors.

**Limitations and Future Directions**

Our study has two notable limitations. Despite recruiting students and community members, we recruited too few community members to analyze that sample separately once discovering that there was a significant difference between the samples. Despite the findings of Bornstein et al. (2017), our findings suggest that researchers should continue to examine both student and community populations.

A further limitation is that we conducted our study online, which could have influenced mock-jurors’ attention or motivation. Even though we included only mock-jurors who passed manipulation checks, it might be useful to conduct future studies in more controlled settings. Another limitation was the strength of the case. The conviction rates in our study reached only 44% even when identifications were correct and obtained using non-suggestive procedures, indicating that the case might not have been strong enough to explore the full
range of the EIC effects. Future research should test the EIC with a stronger case and
manipulate the inclusion of various components of the EIC—the checklist itself might be
responsible for the overbelief. The EIC performed worse than the I-I-Eye (Pawlenko et al.,
2013; Safer et al., 2016); however, we did not directly compare the two novel procedures.
Moreover, the I-I-Eye’s effectiveness was not tested with genuine eyewitnesses, thus, it is
unclear if it can improve mock-jurors’ sensitivity to identification accuracy. Additionally, the
I-I-Eye was examined only in cases with eyewitness testimony rather than in cases that
include both video-recorded identification procedures and eyewitness testimony.
Understanding why (as well as when) the EIC and other traditional legal safeguards are not
effective would provide valuable insights into potential ways of educating jurors about
eyewitness evidence.
Author Contributions

Alena Skalon designed and conducted the study, analyzed the data, and wrote the manuscript. Dr. Jennifer L. Beaudry was closely involved in formulating the research questions, in designing the study, provided assistance with data analysis, and revised the manuscript.

Acknowledgments

This work was partially supported by the Association for Psychological Science Student Grant and Swinburne University of Technology Postgraduate Research Award. The authors would like to thank Travis Edmonds for his assistance with the recording of the Eyewitness Instructions and Checklist.
Supplementary Materials

The Eyewitness Instructions and Checklist: An Alternative Method for Educating Jurors about Eyewitness Evidence

Alena Skalon & Jennifer L. Beaudry

In these supplementary materials, we present the following:

1. Table 18 contains model building statistics predicting students’ and community members’ verdicts as a function of the sample type and its interactions with identification accuracy, identification procedure, educational aid (referenced in the main manuscript on page 16).

2. Table 19 contains hierarchical logistic regression model building summary for students’ verdicts (referenced in the main manuscript on page 17).

3. All results for community participants along with associated tables.

4. Results of Bayesian analyses for both samples along with the associated table.
Table 18

*Binary Logistic Regression Analyses Predicting Mock-jurors’ Verdicts from Sample Type and Identification Accuracy, Identification Procedure, Educational Aid*

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>SE B</th>
<th>Wald</th>
<th>p</th>
<th>OR</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID procedure</td>
<td>0.61</td>
<td>0.53</td>
<td>1.34</td>
<td>.248</td>
<td>1.84 [0.66, 5.16]</td>
</tr>
<tr>
<td>ID accuracy</td>
<td>0.13</td>
<td>0.54</td>
<td>0.06</td>
<td>.808</td>
<td>1.14 [0.40, 3.27]</td>
</tr>
<tr>
<td>Educational aid</td>
<td>0.84</td>
<td>0.53</td>
<td>2.55</td>
<td>.110</td>
<td>2.32 [0.83, 6.51]</td>
</tr>
<tr>
<td>Sample</td>
<td>-0.05</td>
<td>0.46</td>
<td>0.01</td>
<td>.919</td>
<td>0.95 [0.39, 2.36]</td>
</tr>
<tr>
<td>Sample × ID Procedure</td>
<td>-0.27</td>
<td>0.41</td>
<td>0.43</td>
<td>.512</td>
<td>0.76 [0.35, 1.72]</td>
</tr>
<tr>
<td>Sample × ID Accuracy</td>
<td>0.75</td>
<td>0.42</td>
<td>3.21</td>
<td>.073</td>
<td>2.13 [0.94, 4.86]</td>
</tr>
<tr>
<td>Sample × Educational Aid</td>
<td>-0.42</td>
<td>0.41</td>
<td>1.07</td>
<td>.302</td>
<td>0.65 [0.30, 1.47]</td>
</tr>
</tbody>
</table>

*Note. All dfs = 1.*
### Table 19

**Hierarchical Logistic Regression Model Building Summary for Students’ Verdicts**

<table>
<thead>
<tr>
<th>Step</th>
<th>Variable</th>
<th>( \chi^2 )</th>
<th>df</th>
<th>( p )</th>
<th>( \chi^2 )</th>
<th>df</th>
<th>( p )</th>
<th>( B )</th>
<th>( SE )</th>
<th>( OR )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ID procedure</td>
<td>2.28</td>
<td>1</td>
<td>.093</td>
<td>-</td>
<td>-</td>
<td>0.31</td>
<td>0.18</td>
<td>1.37</td>
<td>[0.95, 1.96]</td>
</tr>
<tr>
<td>2</td>
<td>ID accuracy</td>
<td>23.75</td>
<td>2</td>
<td>.001</td>
<td>20.93</td>
<td>.001</td>
<td>0.86</td>
<td>0.19</td>
<td>2.37</td>
<td>[1.63, 3.43]</td>
</tr>
<tr>
<td>3</td>
<td>Educational aid</td>
<td>28.57</td>
<td>3</td>
<td>.001</td>
<td>4.81</td>
<td>.028</td>
<td>0.42</td>
<td>0.19</td>
<td>1.52</td>
<td>[1.05, 2.2]</td>
</tr>
<tr>
<td>4</td>
<td>ID Accuracy × ID Procedure</td>
<td>39.40</td>
<td>4</td>
<td>.001</td>
<td>10.84</td>
<td>.001</td>
<td>-1.30</td>
<td>0.40</td>
<td>0.28</td>
<td>[0.13, 0.61]</td>
</tr>
<tr>
<td>5</td>
<td>ID Procedure × Educational Aid</td>
<td>39.62</td>
<td>5</td>
<td>.001</td>
<td>0.22</td>
<td>.643</td>
<td>0.18</td>
<td>0.38</td>
<td>1.2</td>
<td>[0.57, 2.54]</td>
</tr>
<tr>
<td>6</td>
<td>ID Accuracy × Educational Aid</td>
<td>41.75</td>
<td>6</td>
<td>.001</td>
<td>2.13</td>
<td>.145</td>
<td>0.58</td>
<td>0.40</td>
<td>1.79</td>
<td>[0.82, 3.89]</td>
</tr>
<tr>
<td>7</td>
<td>ID Accuracy × ID Procedure × Educational Aid</td>
<td>43.83</td>
<td>7</td>
<td>.001</td>
<td>2.08</td>
<td>.015</td>
<td>-1.18</td>
<td>0.82</td>
<td>0.31</td>
<td>[0.07, 1.55]</td>
</tr>
</tbody>
</table>

**Final model**: 39.40, 4, .001

*Note.* Final model included ID procedure, ID accuracy, educational aid, ID Accuracy × ID Procedure. For step statistics all df's = 1.
Results for Community Participants

Hypothesis 1: Are Community Members Sensitive to Eyewitness Accuracy?

As expected, community members rendered a guilty verdict more often when they watched correct IDs (51%) compared to mistaken IDs (17%), $B = 1.69$, $SE = .38$, Wald $(1, n = 158) = 19.54, p < .001$, OR $= 5.41 [2.56, 11.43]$ (see Table 20 for model details).

Table 20
Hierarchical Logistic Regression Model Building Summary for Community Members’ Verdicts

<table>
<thead>
<tr>
<th>Step</th>
<th>Variable</th>
<th>Model Statistics for each step</th>
<th>Statistics for each step</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>χ² df p</td>
<td>χ² p B SE B OR</td>
</tr>
<tr>
<td>1</td>
<td>ID procedure</td>
<td>0.03 1 .866</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>ID accuracy</td>
<td>21.37 2 .001</td>
<td>21.35 .001 1.64 .37 5.15 [2.47, 10.72]</td>
</tr>
<tr>
<td>3</td>
<td>Educational aid</td>
<td>21.37 3 .001</td>
<td>0.01 .980 -0.01 .36 0.99 [0.48, 2.03]</td>
</tr>
<tr>
<td>4</td>
<td>ID Accuracy × ID Procedure</td>
<td>22.31 4 .001</td>
<td>0.94 .333 -0.73 0.76 0.48 [0.11, 2.15]</td>
</tr>
<tr>
<td>5</td>
<td>Educational Aid</td>
<td>26.78 5 .001</td>
<td>4.47 .035 -1.58 0.76 0.21 [0.05, 0.92]</td>
</tr>
<tr>
<td>6</td>
<td>ID Accuracy × Educational Aid</td>
<td>26.96 6 .001</td>
<td>0.18 .670 -0.35 0.82 0.71 [0.14, 3.52]</td>
</tr>
<tr>
<td>7</td>
<td>ID Procedure × Educational Aid</td>
<td>29.12 7 .001</td>
<td>2.16 .142 -2.35 1.58 0.10 [0.00, 2.11]</td>
</tr>
</tbody>
</table>

Note. Final model included ID procedure, ID accuracy, educational aid, Educational Aid × ID Procedure. For step statistics all dfs = 1.

Community members who watched correct IDs were less confident in a not guilty verdict, $d = 0.76 [-8.74, 10.25]$, were more likely to believe that the witness correctly identified the culprit, $d = 0.67 [-2.77, 4.11]$, perceived the witness to be more reliable, $d =$
0.59 [0.27, 0.91], and rated the lineup as fairer, \( d = 0.37 [-0.05, 0.80] \), compared to mock-jurors who watched mistaken IDs. ID accuracy did not significantly affect community members’ perceptions of the administrator. For our continuous measures, we report the ANOVA significance tests in Table 21 and the descriptive statistics in Table 22.
Table 21

*Univariate Analyses of Variance Results for ID Procedure × ID Accuracy × Educational Aid Analyses on Continuous Dependent Variables for Community Members (N = 158)*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Confidence-in-verdict</th>
<th>ID likelihood</th>
<th>Lineup fairness</th>
<th>Perceptions of administrator</th>
<th>Perceptions of eyewitness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>$\eta_p^2$</td>
<td>F</td>
<td>$\eta_p^2$</td>
<td>F</td>
</tr>
<tr>
<td>ID accuracy</td>
<td>20.52***</td>
<td>0.12</td>
<td>19.14***</td>
<td>0.11</td>
<td>4.25*</td>
</tr>
<tr>
<td>ID procedure</td>
<td>0.00</td>
<td>0.00</td>
<td>0.36</td>
<td>0.00</td>
<td>17.05***</td>
</tr>
<tr>
<td>Educational aid</td>
<td>0.21</td>
<td>0.00</td>
<td>1.26</td>
<td>0.01</td>
<td>3.83</td>
</tr>
<tr>
<td>ID Procedure × ID Accuracy</td>
<td>1.20</td>
<td>0.01</td>
<td>8.74**</td>
<td>0.06</td>
<td>0.99</td>
</tr>
<tr>
<td>ID Procedure × Educational Aid</td>
<td>5.00</td>
<td>0.03</td>
<td>5.13*</td>
<td>0.03</td>
<td>0.01</td>
</tr>
<tr>
<td>ID Accuracy × Educational Aid</td>
<td>0.05</td>
<td>0.00</td>
<td>0.36</td>
<td>0.00</td>
<td>0.06</td>
</tr>
<tr>
<td>ID Procedure × ID Accuracy × Educational Aid</td>
<td>1.38</td>
<td>0.01</td>
<td>0.89</td>
<td>0.01</td>
<td>0.19</td>
</tr>
</tbody>
</table>

*Note. All dfs = 1.*

***$p < .001$. **$p < .01$. *$p < .05$.  

190
Table 22

Community Members' Mean Ratings (with Standard Deviations) for the Continuous Dependent Variables as a Function of Educational Aid, Identification Accuracy, and Identification Procedure

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th></th>
<th>EIC</th>
<th></th>
<th>Overall</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Correct ID</td>
<td>Mistaken ID</td>
<td>Total</td>
<td>Correct ID</td>
<td>Mistaken ID</td>
<td>Total</td>
</tr>
<tr>
<td>Confidence-in-verdict</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-suggestive</td>
<td>-2.20 (70.08)</td>
<td>-48.25 (48.13)</td>
<td>-22.67 (64.78)</td>
<td>24.71 (63.83)</td>
<td>-40.05 (52.18)</td>
<td>-8.42 (66.13)</td>
</tr>
<tr>
<td>Suggestive</td>
<td>20.05 (67.06)</td>
<td>-27.54 (62.00)</td>
<td>-5.33 (68.05)</td>
<td>-20.21 (59.87)</td>
<td>-40.25 (59.60)</td>
<td>-32.00 (59.64)</td>
</tr>
<tr>
<td>Total</td>
<td>9.20 (68.61)</td>
<td>-35.83 (57.13)</td>
<td>-13.04 (66.77)</td>
<td>6.74 (65.31)</td>
<td>-40.14 (55.14)</td>
<td>-18.83 (64.04)</td>
</tr>
<tr>
<td>ID likelihood</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-suggestive</td>
<td>65.15 (19.66)</td>
<td>34.38 (19.94)</td>
<td>51.48 (24.92)</td>
<td>63.77 (23.46)</td>
<td>43.69 (22.25)</td>
<td>53.49 (24.76)</td>
</tr>
<tr>
<td>Suggestive</td>
<td>61.58 (20.34)</td>
<td>57.84 (22.37)</td>
<td>59.58 (21.29)</td>
<td>51.00 (21.05)</td>
<td>44.90 (21.91)</td>
<td>47.42 (21.45)</td>
</tr>
<tr>
<td>Total</td>
<td>63.32 (19.85)</td>
<td>48.45 (24.16)</td>
<td>55.98 (23.18)</td>
<td>58.66 (23.1)</td>
<td>44.27 (21.83)</td>
<td>50.81 (23.40)</td>
</tr>
<tr>
<td>Lineup fairness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-suggestive</td>
<td>9.30 (1.27)</td>
<td>7.94 (2.60)</td>
<td>8.70 (2.06)</td>
<td>8.43 (2.47)</td>
<td>7.23 (2.47)</td>
<td>7.82 (2.51)</td>
</tr>
<tr>
<td>Suggestive</td>
<td>7.00 (2.69)</td>
<td>6.84 (3.11)</td>
<td>6.92 (2.89)</td>
<td>6.43 (2.93)</td>
<td>5.70 (2.89)</td>
<td>6.00 (2.89)</td>
</tr>
<tr>
<td>Total</td>
<td>8.13 (2.39)</td>
<td>7.28 (2.93)</td>
<td>7.71 (2.69)</td>
<td>7.63 (2.80)</td>
<td>6.50 (2.76)</td>
<td>7.02 (2.82)</td>
</tr>
<tr>
<td>Perceptions of the administrator</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-suggestive</td>
<td>2.48 (1.83)</td>
<td>3.10 (2.65)</td>
<td>2.75 (2.22)</td>
<td>2.10 (1.55)</td>
<td>3.20 (2.13)</td>
<td>2.65 (1.92)</td>
</tr>
<tr>
<td>Suggestive</td>
<td>3.56 (2.58)</td>
<td>3.77 (2.33)</td>
<td>3.67 (2.42)</td>
<td>4.49 (2.62)</td>
<td>4.80 (2.10)</td>
<td>4.67 (2.30)</td>
</tr>
<tr>
<td>Total</td>
<td>3.04 (2.28)</td>
<td>3.5 (2.45)</td>
<td>3.26 (2.36)</td>
<td>3.05 (2.33)</td>
<td>3.98 (2.24)</td>
<td>3.55 (2.32)</td>
</tr>
<tr>
<td>Perceptions of the eyewitness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-suggestive</td>
<td>6.10 (2.04)</td>
<td>4.04 (1.91)</td>
<td>5.19 (2.21)</td>
<td>7.15 (2.16)</td>
<td>5.49 (2.34)</td>
<td>6.30 (2.38)</td>
</tr>
<tr>
<td>Suggestive</td>
<td>6.35 (1.81)</td>
<td>5.99 (1.94)</td>
<td>6.16 (1.86)</td>
<td>6.52 (2.29)</td>
<td>5.38 (1.63)</td>
<td>5.85 (1.98)</td>
</tr>
<tr>
<td>Total</td>
<td>6.23 (1.90)</td>
<td>5.21 (2.13)</td>
<td>5.73 (2.07)</td>
<td>6.89 (2.20)</td>
<td>5.44 (2.01)</td>
<td>6.10 (2.21)</td>
</tr>
</tbody>
</table>

Note. Standard deviations are reported in parentheses.
Hypothesis 2: Does the Suggestiveness of Identification Procedures Affect Community Members’ Sensitivity to Identification Accuracy?

We hypothesized that mock-jurors will be sensitive to ID accuracy when they viewed non-suggestive ID procedures, but that sensitivity will be lost when they viewed suggestive ID procedures. A significant ID Procedure × ID Accuracy interaction emerged only for community members’ ID likelihood ratings. In the non-suggestive condition, community members who watched correct IDs were more likely to believe that the witness correctly identified the culprit, $t(77) = 5.10, p < .001, d = 1.16 [-3.51, 5.84]$, compared to those who watched mistaken IDs. This sensitivity was lost when ID procedures were suggestive, $t(77) = 1.08, p = .284$. The ID Procedure × ID Accuracy did not affect any other variables for the community members sample. Thus, Hypothesis 2 was supported only for one measure.

Hypotheses 3 and 4: Does the EIC Sensitize Community Members to Identification Conditions or to Identification Accuracy?

We predicted that mock-jurors who completed the Eyewitness Instructions and Checklist (EIC) will be more sensitive to identification conditions and/or more sensitive to identification accuracy than those in the control condition. Although a significant Educational Aid × ID Procedure interaction emerged for verdict, $B = -1.52, SE = .76$, Wald $(1, n = 158) = 4.05, p = .044$, OR $= 0.22 [0.05, 0.96]$, the follow-up chi-square tests were not significant, $\chi^2$s $< 2.38, ps > .123$.

A main effect of educational aid did not affect any continuous variables. However, an Educational Aid × ID Procedure interaction affected community members’ confidence-in-verdict, ID likelihood, and perceptions of the eyewitness. Interestingly, the only follow-up test that was significant indicated that community members in the control condition perceived the witness as more reliable when ID conditions were suggestive than non-suggestive, $t(79) =$
2.15, \( p = .034, d = 0.49 \) [0.06, 0.93]. All other follow-up tests were non-significant both for the control, \( ts < 1.58, ps > .118 \), and for the EIC conditions, \( ts < 1.62, ps > .105 \). Overall, neither Hypotheses 3 nor 4 were supported.

Hypothesis 5. Does the EIC Improve Community Members’ Knowledge of Eyewitness Factors?

We expected that mock-jurors who completed the EIC would demonstrate better knowledge of eyewitness factors than mock-jurors in the control condition. As shown in Table 23, the EIC improved community members’ knowledge for three factors: stages of memory (#1), lineup composition (#8), and the claim that photoboards should be banned (#11). Unlike the student participants, the community members demonstrated a good understanding of other factors regardless of the presence of the EIC.
Table 23
Percentage of Correct Answers to Knowledge Questions by Control and the EIC Conditions (Community Members Sample)

<table>
<thead>
<tr>
<th>Question</th>
<th>Correct answer</th>
<th>Control</th>
<th>EIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. What are the stages of memory that are involved in the process of remembering? &lt;select from multiple options&gt;</td>
<td>Encoding; Retention; Retrieval</td>
<td>59%</td>
<td>90%***</td>
</tr>
<tr>
<td>2. A witness may think that s/he is accurate and still be mistaken.</td>
<td>True</td>
<td>98%</td>
<td>100%</td>
</tr>
<tr>
<td>3. Stress increases the accuracy of a witness’ identification.</td>
<td>False</td>
<td>75%</td>
<td>81%</td>
</tr>
<tr>
<td>4. Which of these factors should not be considered when evaluating the time of the crime? &lt;select from multiple options&gt;</td>
<td>Gender of an eyewitness</td>
<td>86%</td>
<td>94%</td>
</tr>
<tr>
<td>5. Presence of a weapon: &lt;select from multiple options&gt;</td>
<td>Decreases the likelihood that the witness will make an accurate identification</td>
<td>61%</td>
<td>71%</td>
</tr>
<tr>
<td>6. Choose the most accurate statement: &lt;select from multiple options&gt;</td>
<td>Eyewitness’s memories can change because of discussions with other people</td>
<td>82%</td>
<td>87%</td>
</tr>
<tr>
<td>7. Which of these factors should not be considered when evaluating the identification procedure? &lt;select from multiple options&gt;</td>
<td>Whether the eyewitness was tired when the photoboard was conducted</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>8. A photoboard in which the suspect stands out is more reliable because it is easier to choose the suspect</td>
<td>False</td>
<td>51%</td>
<td>75%***</td>
</tr>
<tr>
<td>9. Post-identification feedback improve the reliability of eyewitness testimony</td>
<td>False</td>
<td>78%</td>
<td>83%</td>
</tr>
<tr>
<td>10. Post-identification feedback affects: &lt;select from multiple options&gt;</td>
<td>Confidence in the identification decision</td>
<td>61%</td>
<td>68%</td>
</tr>
<tr>
<td>11. Photobords should be banned because they are utterly suggestive</td>
<td>False</td>
<td>78%</td>
<td>91%*</td>
</tr>
</tbody>
</table>

*p < .05. **p < .01. ***p < .001.
Bayesian Analyses: Students and Community Members

We conducted Bayesian ANOVAs for our continuous measures using JASP (JASP Team, 2018) to determine if they support the findings from the inferential statistics. We did not analyze the verdict measure using Bayesian approach because JASP does not support Bayesian binary logistic regression. A Bayesian analysis compares and quantifies the evidence in favor of both the null and the alternative hypotheses (e.g., Wagenmakers et al., 2018; Jarosz & Wiley, 2014). We report inverse Bayes factors (BF\textsubscript{10}), which represent the amount of support for the alternative hypothesis compared to the null hypothesis. In terms of the strength of the evidence, a BF\textsubscript{10} of 1–3 provides weak support for the alternative hypothesis, a BF\textsubscript{10} of 20–150 provides strong support for the alternative hypothesis, and a BF\textsubscript{10} greater than 150 provides very strong support for an alternative hypothesis (Raftery, 1995). Table 24 presents the best models and associated Bayes Factors for each of the continuous dependent variables split by sample.

Hypothesis 1: Are Mock-jurors Sensitive to Eyewitness Accuracy?

Students. According to the inferential statistics, a main effect of ID accuracy affected students’ confidence-in-verdict, ID likelihood, perceptions of the administrator, and perceptions of the eyewitness, but did not affect lineup fairness. The Bayesian analyses supported the findings of the inferential statistics.

Community Members. According to the inferential statistics, a main effect of ID accuracy emerged for community members’ confidence-in-verdict, ID likelihood, lineup fairness, and perceptions of the eyewitness. This, again, was supported by Bayesian analyses.

Hypothesis 2: Does the Suggestiveness of Identification Procedures Affect Mock-jurors’ Sensitivity to Identification Accuracy?
Students. According to the inferential statistics, the ID Procedure × ID Accuracy emerged for students’ confidence-in-verdict, ID likelihood, and perceptions of the eyewitness. The Bayesian analyses confirmed these findings. The ID Procedure × ID Accuracy interaction outperformed all other models for three out of five measures.

Community Members. The significant result from the inferential tests for community members was also supported—the ID Procedure × ID Accuracy affected only the ID likelihood measure for this sample.

Hypotheses 3 and 4: Does the EIC Sensitize Mock-jurors to Identification Conditions or to Identification Accuracy?

Students. According to the inferential statistics, students’ confidence-in-verdict, ID likelihood, and perceptions of the witness were affected by the main effect of educational aid. Again, these findings were supported by Bayesian analyses. However, the inferential tests also revealed a significant ID Procedure × Educational Aid interaction for students’ ratings of lineup fairness. The Bayesian analysis did not support this result; instead, students’ perceptions of the lineup fairness were best explained by the main effect of ID procedure.

Community members. The inferential statistics indicated that an ID Procedure × Educational Aid interaction influenced community members’ ID likelihood, and perceptions of the witness. The Bayesian analyses did not support this interaction and instead indicated that a main effect of educational aid affected only community members’ ratings of lineup fairness. Smaller Bayes factors with the community sample compared to those observed on a student sample, lead us to the conclusion that there was a lack of statistical power and the findings obtained with the community members are less robust than those obtained on the student sample.
Table 24
Bayesian Models with the Highest Bayes Factor for Students and Community Members

<table>
<thead>
<tr>
<th>Variable</th>
<th>Students</th>
<th>Community members</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model</td>
<td>Model</td>
</tr>
<tr>
<td>Confidence-in-verdict</td>
<td>Three main effects&lt;sup&gt;a&lt;/sup&gt;&amp; ID Procedure × ID Accuracy</td>
<td>Main effect of ID Accuracy</td>
</tr>
<tr>
<td></td>
<td>1.30 × 10&lt;sup&gt;7&lt;/sup&gt;</td>
<td>3,136.99</td>
</tr>
<tr>
<td>ID likelihood</td>
<td>Three main effects&lt;sup&gt;a&lt;/sup&gt;&amp; ID Procedure × ID Accuracy</td>
<td>Two main effects&lt;sup&gt;b&lt;/sup&gt;&amp; ID Procedure × ID Accuracy</td>
</tr>
<tr>
<td></td>
<td>7.31 × 10&lt;sup&gt;7&lt;/sup&gt;</td>
<td>661.77</td>
</tr>
<tr>
<td>Lineup fairness</td>
<td>ID Procedure</td>
<td>Three main effects&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>4.37 × 10&lt;sup&gt;18&lt;/sup&gt;</td>
<td>382.23</td>
</tr>
<tr>
<td>Perceptions of administrator</td>
<td>ID accuracy &amp; ID procedure</td>
<td>Main effect of ID procedure</td>
</tr>
<tr>
<td></td>
<td>1.05 × 10&lt;sup&gt;14&lt;/sup&gt;</td>
<td>179.93</td>
</tr>
<tr>
<td>Perceptions of the eyewitness</td>
<td>Three main effects&lt;sup&gt;a&lt;/sup&gt;&amp; ID Procedure × ID Accuracy</td>
<td>Main effect of ID accuracy</td>
</tr>
<tr>
<td></td>
<td>8.65 × 10&lt;sup&gt;27&lt;/sup&gt;</td>
<td>78.63</td>
</tr>
</tbody>
</table>

<sup>Note.</sup> BF<sub>10</sub> illustrates the amount of support in favor of the alternative hypothesis over the null hypothesis.
<sup>a</sup>Identification accuracy, identification procedure, and educational aid. <sup>b</sup>Identification accuracy and identification procedure.
Chapter 7. Conclusion

This thesis had four main aims. The first research aim was to establish the effectiveness of traditional legal safeguards—judicial instructions and expert testimony—in sensitising mock-jurors to the quality of identification procedures and to eyewitness identification accuracy (Papers 1 and 2). The second research aim was to evaluate whether the effectiveness of expert testimony can be improved by accompanying it with visual aids (Paper 2, Study 1). The third research aim was to test whether the type of eyewitness evidence (a video-recorded identification procedure and the eyewitness’s testimony or the eyewitness’s testimony alone) affects mock-jurors’ perceptions of eyewitnesses and the effectiveness of expert testimony (Paper 2, Study 2). Finally, the fourth research aim was to examine the effectiveness of an innovative educational aid—the Eyewitness Instructions and Checklist (EIC)—in sensitising mock-jurors to the quality of identification procedures and to identification accuracy. In this conclusion chapter, I summarise the findings of the three empirical papers (Papers 1–3), discuss how they pertain to the research aims, discuss future directions, outline limitations of the research, and consider the practical implications of my work.

Paper 1

In this paper, I explored the effectiveness of the Victorian judicial instructions, which have not been previously tested empirically. This study used a minimal trial paradigm; the eyewitness evidence was the only piece of evidence and there was no case summary or attorneys’ arguments. Participants \( n = 232 \) watched an identification procedure and the testimony of the same witness and were asked to render a verdict and complete a questionnaire. I varied whether participants saw correct or mistaken identifications obtained using non-suggestive or suggestive identification procedures, as well as whether they read the judicial instructions or completed a
filler activity. In line with previous research with judicial instructions from the USA (e.g., Cutler et al, 1990; Jones, Bergold, Dillon, & Penrod, 2017) and New South Wales (Martire & Kemp, 2009), the Victorian judicial instructions did not sensitise participants to identification accuracy or to identification conditions. Although judicial instructions did not affect verdicts, they did affect participants’ estimates of the likelihood that the eyewitness correctly identified the culprit. Contrary to the information contained in the judicial instructions, participants assigned higher likelihood scores when identifications were obtained using suggestive identification procedures compared to non-suggestive identification procedures. In other words, the judicial instructions led to confusion about the influence of suggestive identification procedures on eyewitnesses’ decisions. There was another important finding. Regardless of the presence of judicial instructions, participants were sensitive to identification accuracy, but only when identifications were obtained using non-suggestive identification procedures. In contrast, when identification procedures were suggestive, participants believed correct and mistaken eyewitnesses approximately at the same rate.

**Paper 2**

Paper 2 included two studies. I used the whole-trial paradigm to test the effectiveness of traditional expert testimony and expert testimony with visual aids. In the second study, in addition to testing the effectiveness of expert testimony, I tested whether the type of evidence influences the effectiveness of expert testimony. In Study 1, the eyewitness stimuli were similar to that presented in Paper 1, such that mock-jurors watched a video-recorded identification procedure and the eyewitness’s testimony. In Study 2, I manipulated the type of evidence; mock-jurors either watched the same stimuli as in Study 1 or the eyewitness’s testimony only. In both studies, mock-jurors (study 1, \( N = 561 \); study 2, \( N = 478 \)) watched a case summary delivered by
the judge and the attorneys’ opening statements before they watched the eyewitness evidence.

Following the presentation of eyewitness evidence, mock-jurors either watched traditional expert testimony, expert testimony with visual aids (study 1 only), or completed a filler activity. Mock-jurors then watched the attorneys’ closing arguments, rendered a verdict and completed a questionnaire.

Expert testimony did not improve mock-jurors’ sensitivity to identification procedures or to identification accuracy and instead led to skepticism. Mock-jurors who watched expert testimony estimated a lower likelihood that the witness correctly identified the culprit (cf. control condition) regardless of identification conditions or identification accuracy. Expert testimony with visual aids did not outperform the traditional expert testimony. Similarly to findings described in Paper 1, mock-jurors were sensitive to identification accuracy when identifications were obtained using non-suggestive identification procedures. However, this sensitivity was lost when the identification procedures were suggestive. Additionally, Study 2 demonstrated that mock-jurors who watched the eyewitness’s testimony without the video-recorded identification procedure were more likely to believe that the witness correctly identified the culprit than mock-jurors who watched both.

Paper 3

The final paper included in this thesis tested the effectiveness of an innovative educational aid—the Eyewitness Instructions and Checklist (EIC). The EIC consists of a narrated PowerPoint presentation and a checklist that asks mock-jurors to evaluate the presence of each eyewitness factor in the case at hand. Mock-jurors (N = 536 students; N = 158 community members) watched the EIC (or a control video) before the trial began. They then read the case summary and attorneys’ arguments, and watched the eyewitness evidence, which varied in terms
of suggestiveness of identification procedures and identification accuracy. Following the evidence presentation, mock-jurors read attorneys’ closing statements, completed the second part of the EIC—the checklist (or a control activity)—and completed a questionnaire. Results indicated that the EIC did not improve mock-jurors’ sensitivity to the quality of identification procedures or to identification accuracy and instead it led to overbelief. Mock-jurors who watched the EIC rendered more guilty verdicts than mock-jurors who completed the control activity. In addition, this paper supported the findings of Papers 1 and 2: mock-jurors were sensitive to identification accuracy when identification procedures were non-suggestive, but not when they were suggestive.

**Synthesis of Findings**

**Research aim 1.** My studies demonstrated that traditional legal safeguards are ineffective in sensitising mock-jurors to the quality of identification procedures or to identification accuracy. Neither judicial instructions (Paper 1), nor expert testimony (Paper 2) improved the accuracy of mock-jurors’ assessments of eyewitness evidence. Moreover, judicial instructions led to confusion, and there was some indication that expert testimony led to skepticism. The differences in the content and presentation of judicial instructions and expert testimony might be responsible for the different effects of these legal safeguards. This is discussed in detail in Chapter 2 of this thesis. To recap, judicial instructions are read by the judge at the end of the trial along with all other instructions; moreover, they are usually formulated as a series of questions and do not provide an explanation of eyewitness factors. Expert testimony, on the other hand, is a dialogue between the expert and one of the attorneys (with the potential for cross-examination). The expert provides background information about eyewitness research and usually explains how various factors can affect eyewitnesses.
At this point, researchers have not yet examined the question of whether jurors process the same information provided by the judge and the expert in a different manner. Expert testimony might lead to skepticism because jurors rely on the expert’s credentials rather than carefully processing the content of the expert’s testimony. According to the elaboration likelihood model, people use either the central or the peripheral information-processing route (Petty & Cacioppo, 1986). If a central route is used, a person carefully scrutinises the content of the message, hence, the quality of the content determines the persuasiveness. In contrast, if a person uses the peripheral route, he or she does not process all relevant information and instead relies on other cues, such as the status of the person who delivered the message. According to the elaboration likelihood model, people are more likely to engage in effortful decision-making and carefully consider information—that is, use the central route—when they are motivated and able to do so. When the material is complex, the level of motivation must be high to ensure a person’s reliance on facts rather than stereotypes and other peripheral information. Empirical evidence in other domains has demonstrated that jurors do in fact use peripheral information, such as the expert’s level of expertise when the content of the testimony is complex, but not when it is simple (Cooper et al., 1996; Cooper & Neuhaus, 2000). When evaluating such complex information as that provided in expert testimony or judicial instructions, jurors may use the credentials of the judge and the expert to decide how to utilise the information (Bank & Poythress, 1982). The fact that eyewitness experts are usually hired by the defence might skew jurors’ perceptions of the eyewitness evidence, which would explain the skepticism effect.

**Research aim 2.** In an attempt to improve the effectiveness of expert testimony, I added visual aids to accompany the expert testimony. In contrast to research in other domains (e.g., Jones & Kovera, 2015; Park & Feigenson, 2013), Paper 2, Study 1 demonstrated that visual aids
did not improve the effectiveness of expert testimony. However, this provides further support to
findings outlined in an unpublished doctoral thesis in which Binder (2006) used PowerPoint
slides illustrating key points of expert testimony on the psychology of eyewitness identification
and found no effect of visual aids on participants’ verdicts. Supplementing expert testimony with
visual aids might not be effective if expert testimony’s ineffectiveness is not due to poor
comprehension or lack of memory for the information. Skepticism consistently caused by expert
testimony might be rooted in some unique characteristics of expert testimony—these
characteristics are yet to be determined.

**Research aim 3.** In addition to testing the effectiveness of legal safeguards, I explored
whether the type of eyewitness evidence affects mock-jurors’ perceptions of the evidence and of
expert testimony. A number of scholars recommended video-recording the initial identification
procedure (e.g., Kassin, 1998; National Research Council, 2014). A video-record preserves
information about the eyewitness’s and the administrator’s behaviour at the time of the
identification. Once these videos are created, they are likely to be admitted in court as part of the
eyewitness evidence. This poses significant challenges, however, because researchers are only
beginning to explore jurors’ perceptions of video-recorded identification procedures (i.e.,
Beaudry et al., 2015; Modjadidi & Kovera, 2015, Reardon & Fisher, 2011). No studies, to my
knowledge, manipulated the type of evidence, the suggestiveness of identification procedures,
and identification accuracy, along with expert testimony.

Paper 2, Study 2 demonstrated that mock-jurors who watched the eyewitness’s testimony
alone were more willing to believe eyewitnesses than those who watched the video-recorded
identification procedure along with the eyewitness’s testimony. This finding indicates that mock-
jurors’ tendency to overbelieve eyewitnesses might be reduced if they watch the video-recorded
identification procedure in addition to the eyewitness’s testimony; however, it is too early to make any practical recommendations based on this result. Moreover, the type of evidence also affected mock-jurors’ perceptions of eyewitness evidence depending on the presence of expert testimony. Mock-jurors in the control condition were sensitive to identification accuracy when they watched both a video-recorded identification procedure and the eyewitness’s testimony (in line with Reardon & Fisher, 2011), whereas those in the expert testimony condition were sensitive to identification accuracy when they watched the eyewitness’s testimony only. This finding suggests that video-recorded identification procedures are a fundamentally different type of evidence (cf. eyewitness testimony) and future research should establish conditions under which it should (or should not) appear in front of the jury.

**Research aim 4.** Taken together, the results of empirical findings suggest that a new educational aid is needed. For example, Mnookin (2015) proposed an alternative method of educating the jury—a made-in-advance modular expert testimony. She argued that expert testimony is a more appropriate approach to educating the jury than judicial instructions because “[judicial instructions] are in a way, a category mistake: an effort to turn something that is fundamentally evidence into an appropriate topic for judicial instruction” (p. 1815). However, experts are extremely costly and are limited in supply; thus, they are not available in the vast majority of cases that involve eyewitness evidence. Her proposed solution is creating expert testimony modules about various eyewitness factors in which the expert would testify in response to questioning by the defence attorney and by the prosecutor. This approach would draw on some of the strengths of expert testimony (e.g., presenting information in response to questioning; cross-examination) while ensuring that it is cost-effective and is readily available.
This idea should be tested empirically; however, the demonstrated ineffectiveness of expert testimony suggests that more novel solutions are needed.

Based on these considerations, I created the Eyewitness Instructions and Checklist (EIC) that addressed limitations of traditional legal safeguards. Both judicial instructions and expert testimony are presented after the evidence, rather than before the evidence. By the time jurors receive guidance on how to evaluate eyewitness evidence, they might not be able to retrospectively notice any issues with the evidence. Furthermore, empirical tests of these two educational aids cannot necessarily be extrapolated to real-life trials. Both judicial instructions and expert testimony suffer from a lack of uniformity. Judges can usually modify the content of judicial instructions and experts follow the same general structure, but can still present information differently. This is problematic because any modifications to judicial instructions or expert testimony that were empirically tested would make the effectiveness of these methods unknown (although it is unlikely that any modification will be more effective than that established in experiments). The only solution to this issue would be the creation of a standardised educational aid. A standardised aid would ensure that conclusions reached in empirical research are transferable to other situations because the content and the presentation of the aid will remain unchanged.

Despite a promising theoretical rationale and similarities with an educational aid that was shown to be effective (i.e., the interview-identification-eyewitness [I-I-Eye] educational aid), the EIC did not improve mock-jurors’ sensitivity to identification conditions or to identification accuracy and instead led to overbelief. The EIC led to overbelief even when the identification was suggestive and/or mistaken. This is in contrast to nearly all previous research on this topic. Only one study, to my knowledge, demonstrated overbelief as the result of an educational aid
(Ramirez, Zemba, & Geiselman, 1996). They manipulated the timing of the presentation of judicial instructions and found that if the instructions were presented before the evidence, mock-jurors’ sensitivity to witnessing and identification conditions was preserved, whereas if the instructions were presented after the evidence, they led to skepticism. Presentation both before and after the evidence is similar to the presentation of the EIC in my study. The first component of the EIC was presented before the evidence, whereas the second component was presented after the evidence. Due to the very limited number of studies conducted to date, it would be premature to make any definite conclusions regarding the effect of timing on jurors’ judgments.

Another potential explanation for mock-jurors’ overbelief is the presence of the checklist. The checklist was designed to help mock-jurors integrate the information provided in the first part of the EIC to the case at hand. However, it might have caused overbelief by drawing attention to the eyewitness factors that were good in the case, and distracting from poor factors. Both possibilities warrant further exploration.

**Additional findings.** Even though not stated as one of my research aims, one of the key contributions of this thesis was the consistent demonstration of the damaging effects of suggestive identification procedures on mock-jurors’ ability to differentiate between correct and mistaken identifications. Suggestive identification procedures impaired mock-jurors’ sensitivity to identification accuracy in all four studies. This effect occurred even though the suggestive identification procedures in Papers 1 and 3 were different from those in Paper 2. In Papers 1 and 3, the suggestive identification procedures were: single-blind administration, suggestive lineup instructions, and a biased lineup. In Paper 2, the suggestive procedures were: single-blind administration and the presence of post-identification feedback.
Lack of a Theoretical Model

The mixed results obtained in studies that tested the effectiveness of various educational aids (both in my thesis and in the literature) highlight that even though research has established the most consistent effects of judicial instructions and expert testimony, and conducted evaluations of alternative educational aids, there is still little understanding of why some educational aids are more persuasive or effective than others. A theoretical model would help explain the results, identify components of educational aids that are effective, as well as determine circumstances in which they can be effective. A model would also support the integration of research that examined the effectiveness of legal safeguards in eyewitness cases with the broader jury decision-making literature. Studies of eyewitness experts and judicial instructions on eyewitness issues have largely ignored characteristics of the defendant (e.g., physical attractiveness, socio-economic status, demeanour), of the expert (e.g., gender, credentials), of the jurors (e.g., socioeconomic status, need for cognition, extraversion; for an overview see Devine, 2012), and of eyewitnesses (e.g., bystander vs. victim). Without knowing how educational aids interact with various other factors present at a trial, our understanding of the effectiveness of an educational aid will be incomplete. It is also unclear how educational aids affect story construction. According to the story model of jury decision-making (Pennington & Hastie, 1986), jurors construct stories about the events in question based on the facts of the case as presented during the trial. The ability of an educational aid to complement or disrupt the story has, to my knowledge, never been examined.

In order for the field to progress further, an integrative model of jury decision-making should be created. Devine (2012) proposed a multi-level theory that includes two models. The “director’s cut” model incorporates empirically-supported factors that influence jurors’ pre-
deliberations decisions (e.g., pre-trial publicity, socio-economic status of the defendant, jurors’ need for cognition). The “story sampling” model of jury decision-making explains the interaction between factors that affect jury verdict (e.g., deliberation style, timing of polling). Eyewitness evidence might constitute one of the factors at the pre-deliberation decision-making stage by, for example, strengthening a story offered by the defence. It is unclear, however, where an educational aid can fit in Devine’s model or whether it needs to be a part of that specific model at all. It is possible that an educational aid would act as a mediator to one of the other characteristics (e.g., the eyewitness’s trustworthiness) and, thus, affect the story construction.

Importantly, the theoretical framework should be created based on findings of studies that used a genuine eyewitness paradigm (Wells, Lindsay, & Ferguson, 1979). This is the only paradigm that can answer the ultimate question—whether educational aids help jurors make more accurate judgments, believing correct identifications and discounting mistaken ones (Devine, 2012; Martire & Kemp, 2011). This paradigm comes closest to replicating real jurors’ task of evaluating the eyewitness evidence in its entirety. Mock-jurors who watch genuine eyewitness evidence can assess the eyewitness’s demeanour (e.g., eyewitness’s confidence, verbal and non-verbal cues), as well as the content of their testimony (e.g., reported witnessing and identification conditions). To date, a very limited number of studies used a genuine eyewitness paradigm to examine the effectiveness of educational aids (i.e., Martire & Kemp, 2009; Wells, Lindsay, & Tousignant, 1980; all papers presented in this thesis).

One of the key reasons for using a genuine eyewitness paradigm is that simulations do not permit the examination of the interaction between the identification procedure and identification accuracy. Simulations use actors who did not witness a crime and did not make an identification. Actors can recreate some behavioural cues of genuine eyewitnesses (e.g., confidence) and report
varying witnessing and identification conditions but they cannot recreate the full spectrum of an eyewitness’s behaviour. The damaging influence of suggestive identifications procedures on mock-jurors’ ability to differentiate between correct and mistaken eyewitnesses replicated previous work (Beaudry et al., 2015; Smalarz & Wells, 2014) and is one of the most consistent findings of this thesis. Findings of this thesis highlight that suggestive identification procedures irreversibly damage mock-jurors’ perceptions of eyewitness evidence and none of the examined educational aids restored mock-jurors’ sensitivity to identification accuracy. This suggests that it might be more appropriate to exclude identifications obtained using suggestive identification procedures rather than attempt to educate the jury about dangers of suggestive identification procedures, at least until an effective educational aid is created. An additional benefit of the exclusionary approach would be that police will be more likely to follow best practice identification procedures if identifications obtained using suggestive procedures were being excluded.

**Incorporating Psychological Findings into Legal Practice**

Understanding the effectiveness of educational aids is a necessary step towards improving trial practices; however, it is not up to psychologists to decide what will or will not be incorporated into legal practice. For this reason, it is of crucial importance to make sure that legal scholars understand psychological findings and appreciate their importance and, at the same time, that psychologists address issues raised by legal scholars. In this section, I consider factors that can affect the criminal justice system’s willingness to incorporate psychological research findings into trial practices.

Jurors’ decision-making studies attempt to simulate in-court experiences and some critics point out that, due to their methodological limitations, conclusions reached in these studies
should be regarded with special caution and cannot be extrapolated to the processes happening in real trials. Leverick (2016) described several common limitations of juror decision-making research: the use of students as participants, presenting a trial as a transcript, the absence of deliberation, and the fact that participants are not invested in the outcome of the mock-trial, because they are aware that this is a simulation and not a real case.

The use of students as participants in the vast majority of juror decision-making studies is of concern to legal scholars because students are not representative of the general population (as is the case with jurors). A meta-analysis of 53 jury studies (17,716 participants) indicated that there were no significant differences in the performance of student and non-student participants when it comes to verdicts, culpability decisions, or sentencing judgments (Bornstein et al., 2017). It is worth noting, however, that I found significant differences between students and community members (described in Paper 3). Jones et al. (2017) also compared students and community members and discovered that, even though the two samples did not differ in their verdicts, community members were more likely to believe that the eyewitness correctly identified the culprit. These differences are in contrast to the results of the Bornstein et al.’s (2017) meta-analysis and require further investigation.

Leverick (2016) outlined another limitation: providing participants with a transcript instead of a video-recorded trial or a live re-enactment of a trial. Reading a transcript is no doubt a very different experience than sitting in a courtroom. Critically, the majority of studies examining the effectiveness of educational aids used video-taped stimuli (e.g., Cutler, Dexter, & Penrod, 1990; Fox & Walters, 1986; Greene, 1988; Jones et al., 2017; Katzev & Wishart, 1985; Wells et al., 1980). The results of expert testimony studies that used transcripts (Geiselman et al. 2002; Geiselman & Mendez, 2005) are generally in line with the results obtained in studies that
presented a video-recorded trial; thus, it appears that the presentation medium does not significantly affect the results. I utilised either a transcript or a video of a case summary and attorneys’ arguments. Furthermore, I used video-recorded genuine eyewitness evidence and added video-recorded identification procedures—a new type of evidence that might be presented to jurors in real cases.

According to Leverick (2016), another limitation of mock-juror studies is the lack of deliberation. Deliberation is believed to decrease individual stereotypes and clarify the evidence (Diamond, 1997); thus, in the absence of deliberation, she argued that the conclusions of mock-juror studies might not be informative. The evidence, however, shows that deliberation can sometimes exaggerate misconceptions instead of reducing them (Devine, 2012). There is also some evidence that deliberation can correct misconceptions held by the minority if the majority of jurors correctly understood the instructions (Diamond, 1997). It is worth testing how deliberation affects the effectiveness of aids that educate jurors about eyewitness evidence.

Finally, simulated trials that are used in psychological research might reduce the ecological validity of the studies because participants are not as invested in the outcome of the trial as are jurors in real cases. Although this is true, I argue that this investment comes at a cost. In real cases, jurors have to listen to evidence for days or weeks, which increases their cognitive load and reduces their ability to pay attention to information presented during the trial (Kleider-Offutt, Clevinger, & Bond, 2016). Even though participants’ investment might be lower than that of real jurors, in laboratory studies participants can focus on the evidence presented to them in an hour-long trial (e.g., Jones et al., 2017) and pay attention to the information without being cognitively overloaded. Even under these “ideal” conditions, however, judicial instructions have no effect on mock-jurors’ judgments and expert testimony most often leads to skepticism. It is unlikely that in
real cases jurors would incorporate the judge’s instructions or the expert’s information into their evaluation of the evidence more effectively than mock-jurors in the laboratory studies. Thus, the results of mock-juror studies not only provide valid conclusions about the effectiveness of educational aids, they might, in fact, overestimate their effectiveness.

In summary, psychological and legal scholars have a different approach to assessing the ecological validity of the studies. This should be explored further if psychologists want their research to be considered by judicial personnel. The gap between legal practice and psychological research should be narrowed—it is essential that lawyers and psychologists discuss the research methodology and findings, and address issues that are articulated by legal scholars.

**Limitations**

In addition to the limitations discussed in each paper, one of the limitations is that cases used in this research produced moderate conviction rates which may have limited the ability of educational aids to further reduce the conviction rates. In Paper 1, participants relied only on the information provided by the witness and there was no circumstantial evidence, whereas in Papers 2 and 3 mock-jurors watched (Paper 2) or read (Paper 3) a case summary and opening and closing statements that mentioned other factors present in the case (e.g., the defendant’s car matched the car of the culprit). It is possible that participants in my studies were not persuaded by the eyewitnesses. Eyewitnesses testified about the video-recorded crime that they saw, which is a very different experience from witnessing a real crime—this aspect might have influenced eyewitnesses’ demeanour in my studies and, hence, influenced mock-jurors’ perceptions of the evidence. Furthermore, the administrators of the lineup and eyewitnesses were all young people, which potentially detracted from the ecological validity of the studies.
Similarly, any findings of this thesis might be limited to eyewitness evidence used in my studies. The eyewitnesses might have exhibited behaviour that increased mock-jurors’ sensitivity to identification accuracy (e.g., very fast identification for correct eyewitnesses). The differences might be less pronounced with other eyewitnesses. The results of my studies, however, support findings of previous research that used an entirely different set of stimuli and found that mock-jurors were sensitive to identification accuracy when identification procedures were non-suggestive (Reardon & Fisher, 2011), but they could not discriminate between correct and mistaken eyewitnesses when identification procedures were suggestive (Smalarz & Wells, 2014). Moreover, to increase stimulus sampling, I used four eyewitnesses in each of the conditions and two different types of suggestive identification procedures. In addition to single-blind administration in all studies, Papers 1 and 3 used biased lineup and suggestive lineup instructions, whereas Paper 2 used post-identification feedback. The damaging effect of suggestive identification procedures occurred in all four studies despite the differences in identification procedures. Undoubtedly, more eyewitness stimuli should be created to ensure that findings of empirical research are not limited to a particular set of stimuli and can be extrapolated to real-life situations. Future research should create additional eyewitness stimuli using eyewitnesses of various ages, who witnessed various types of crime, and employ actors to play the administrators of the lineup who would be of similar age to police staff.

Future Research

As discussed above, expert testimony is more persuasive, but not more effective than judicial instructions, even though there is no explanation of why this is the case. Future research should examine which components of expert testimony make it more persuasive than judicial instructions and build a new educational aid using this information. For example, how the
information is presented is as important as *what* information is presented. Tests of the Henderson instructions demonstrated that even the most up-to-date and empirically-informed instructions do not influence mock-jurors’ judgments (e.g., Jones et al., 2017). Future work should focus on determining the underlying mechanisms and building a theoretical model. Without a theoretical model, our understanding of the effectiveness of educational aids is limited to stating mixed results. One of the first steps towards building a model would be the systematic manipulation of various components of educational aids, as well as systematic manipulation of witnessing conditions, identification conditions, and identification accuracy. For example, future research should examine whether the expert’s credentials affect jurors’ perceptions and how that varies across various factors of the eyewitness evidence.

Furthermore, witnessing and identification conditions should be manipulated independently. There is some evidence that jurors are sensitive to some identification conditions (e.g., foil bias), but not others (e.g., instruction bias; Devenport et al., 2002). Empirical research should establish which identification conditions are most problematic to jurors and focus on improving jurors’ sensitivity to those specific conditions. Moreover, a more recent study that independently manipulated witnessing conditions and identification conditions found that mock-jurors are sensitive to the latter, but not to the former, regardless of the presence of a legal safeguard (Jones et al., 2017). Without a systematic manipulation it is unclear if this finding is the result of some meaningful differences or if it occurred by chance.

One common limitation of educational aids on eyewitness evidence is that they focus on discussing the quality of witnessing and identification conditions, which is associated with the probability of making a correct identification. However, they do not provide any assistance in evaluating the *accuracy* of identifications. Setting aside the issue of whether such assessment
would be permitted in court, eyewitness literature provides very limited guidance on factors that can be indicative of identification accuracy (Kaminski & Sporer, 2017; Weber, Brewer, Wells, Semmler, & Keast, 2004; Wixted & Wells, 2017) and no guidance at all on how to educate mock-jurors about those factors. In addition, the reliability of emerging types of eyewitness evidence (e.g., Facebook identifications, see Appendix 1), as well as jurors’ perceptions of these types of evidence should be empirically tested. Researchers should provide police and lawyers with empirically-supported guidance on handling such evidence. If, for example, Facebook identifications are admitted in court, any new educational aid should also explain the psychology of this new type of identification evidence.

Finally, the consistent damaging effects of suggestive identification procedures on mock-jurors’ ability to differentiate between correct and mistaken identifications across my four studies highlight the need for future research to examine how the interaction of identification procedures and identification accuracy might affect real-life jurors. For example, if the underlying issue with suggestive identification procedures is that they artificially increase eyewitnesses’ confidence and make eyewitnesses appear more persuasive to mock-jurors, then it is worth noting that in real trials eyewitnesses are more likely to appear confident because they went through pre-trial preparation, were exposed to pre-trial publicity, etc. (see Wixted & Wells, 2017). The effectiveness of educational aids should be examined with eyewitnesses who went through pre-trial preparation.

The damaging effect of suggestive identification procedures raises another question: what should the aim of educational aids be? Is it to educate the jury about (1) factors that can affect identification accuracy; (2) factors that are indicative of identification accuracy; or (3) suggestive identification procedures, which might help jurors restore their initial sensitivity to eyewitness
identification? It is essential that the consensus about the ultimate aim of educational aids is established among psychological researchers, as well as between psychologists and lawyers. Knowing the aim would inform the most appropriate methodology for conducting empirical research, as well as the content of any new educational aids.

Conclusion

This thesis includes studies that were the first, to my knowledge, to test the effectiveness of educational aids in sensitising mock-jurors to eyewitness evidence presented via a video-recorded identification procedure and the eyewitness’s testimony. Overall the findings of this thesis are consistent with that of previous research: traditional educational aids (i.e., judicial instructions and expert testimony) are not effective in sensitising mock-jurors to identification conditions or to identification accuracy. This further highlights the need to reassess traditional approaches to educating jurors about eyewitness evidence. I created the Eyewitness Instructions and Checklist educational aid to address the limitations of traditional educational aids. Despite the promising rationale, it did not sensitise mock-jurors either. Although this result is contrary to my predictions, it nevertheless advances our knowledge by demonstrating that this method of educating the jury is not effective. Future research should establish why the Eyewitness Instructions and Checklist was not effective, but I would argue that the checklist might not be suitable for the task of sensitising jurors to the quality of identification conditions or to identification accuracy. Despite not having answers to all questions, psychological knowledge grows every day, today we know more about eyewitness evidence and juror decision-making than we did 20 years ago. In order to further advance the field, psychologists and legal scholars should combine their efforts and design studies together, write book chapters that combine...
knowledge from both fields, and participate in joint conferences, contributing their expertise, openly sharing their concerns, and reaching a consensus about solutions.
References


Devenport, J. L, Kimbrough, C. D, & Cutler, B. L. (2009). Effectiveness of traditional safeguards against erroneous conviction arising from mistaken eyewitness identification. In


Evidence Act 1995 (NSW) Retrieved from


Gittany v. R NSWCA 182 (2016)


*IMM v. The Queen* HCA 14 (2016).


*Manson v Braithwaite* 432 US 98 (1977).


*Newsome v. McCabe* 319 F. 3d 301 (2003).


doi:10.1111/lcrp.12048


doi:10.1002/acp.2895


R v. McIntosh 35 OR (3d) 97, 177 CCC (3d) 385 (CA) (1997).


R. v. Tat CanLII 2234 (ON CA) (1997).


*United States v. Telfaire* 469 D.C. Cir. 552 (1972).


The Science behind Bayley v The Queen (2016)

Alena Skalon & Jennifer L. Beaudry

Swinburne University of Technology

Author note: Alena Skalon, Department of Psychological Sciences; School of Health Sciences; Faculty of Health, Arts and Design; Swinburne University of Technology, Hawthorn, Victoria, Australia; postal address: PO Box 218, Hawthorn, Victoria 3122 Australia; telephone: +61 3 9214 5028; email: askalon@swin.edu.au.

Jennifer L. Beaudry, Department of Psychological Sciences; School of Health Sciences; Faculty of Health, Arts and Design; Swinburne University of Technology, Hawthorn, Victoria, Australia; postal address: Mail H24, PO Box 218, Hawthorn, Victoria 3122 Australia; telephone: +61 3 9214 5028; email: jbeaudry@swin.edu.au.

Please note that this is a pre-print version of the article.

Received 01 Feb 2017, Accepted 21 Jul 2018, Published online: 30 Jan 2019
Abstract

Eyewitness identification was the principal evidence in *DPP v Bayley* (2015), in which Adrian Bayley received a jail sentence for the rape and assault of a woman that took place in 2000. Twelve years after the attack, the victim identified Bayley from a photograph seen on Facebook and later in a formal police identification procedure. At the time of the initial Facebook identification the victim knew about Bayley’s involvement in Gillian Meagher’s case. Bayley successfully appealed this conviction in 2016. The Court of Appeal held that the identification evidence in this case had multiple weaknesses and should not have been permitted at the initial trial. In the decision, the Court relied on legal precedents to support their judgement. In this article we review the empirical evidence regarding each of the issues raised by the Court. In addition, we review how the stressfulness of an event can influence the reliability of an identification, and explore why the jury may have rendered a guilty verdict based on the weak identification evidence.

*Keywords:* eyewitness identification, identification procedure, photoboard identification, identification evidence, Facebook identification, Bayley, eyewitness evidence
The Science behind *Bayley v The Queen* (2016)

Identification evidence often plays an important role in criminal trials and *DPP v Bayley* (2015; conviction trial 1) is one of the examples in which a defendant was convicted based almost entirely on an identification made by the victim. Alarmingly, the jury convicted Adrian Bayley despite multiple weaknesses in the identification evidence in this case. The Court of Appeal ruled that the evidence should not have been permitted in the initial trial due to a number of factors. In this article, we review the empirical evidence regarding each of the issues raised by the Court. Specifically, we discuss how delay can affect identification accuracy, potential problems with an identification from a single photograph, the influence of expectations on people’s decisions, and how a change in the perpetrator’s appearance between the crime and identification can affect identification accuracy. We also discuss how prior exposure to the suspect’s photograph can influence subsequent identification attempts (referred to by the courts as the displacement effect), and discuss the distinction between evidence of identification and evidence of similarity. We also raise two additional factors not mentioned in the appeal decision: the impact of stress on the reliability of an identification, and potential explanations for why the jury convicted Adrian Bayley at the initial trial.

This case brings to life some of the most-discussed issues associated with eyewitness identification evidence. Who should determine the reliability of this evidence? When should it be deemed inadmissible? Can juries appropriately evaluate the quality of identification evidence? Psychological science can provide some insight into the best ways to collect and handle identification evidence, as well as the ways of presenting this evidence to the jury to ensure the defendant’s right to a fair trial.

**The Facts of the Case**
In November 2000, an 18-year old female (referred to as GH) was walking in St Kilda when a Mini Minor car approached her and the driver suggested paying money in exchange for sex. GH agreed and got into the car. Shortly after they drove away, the driver punched GH in the face, drove down a narrow alley, blocked the passenger’s door, and raped her. GH eventually escaped from her attacker. She testified that she decided not to report the crime to the police because she thought that the police would not believe her. She did, however, report the incident to a health service for those involved with street-based sex work, and placed a description of her attacker on a noticeboard. According to her description, the man had “short, blonde spikey hair,” “ranga arms,” and was “evil eyed during the attack” (Bayley v The Queen, 2016, at [26]).

Twelve years after the attack, GH saw a photo of Adrian Bayley on Facebook when she was scrolling through a missing persons page. At that time, she knew that Bayley had been arrested on 27 September for the rape and murder of Gillian Meagher (Bayley v The Queen, 2016). She called Crime Stoppers to report that she recognised him as the man who raped her in 2000. The photo from which GH identified Bayley had been taken 11 years after her rape. On 29 October 2012, she made a statement to a member of the Homicide squad of the Victoria Police. On 1 February, 2013, GH was informed that Adrian Bayley had been charged with multiple offences against her, and 10 days later she was invited to the police station to view a photoboard of 12 men. GH admitted that between the initial Facebook identification and the photoboard identification, she saw Bayley’s photo in the media on multiple occasions in relation to Gillian Meagher’s case. When GH identified the photo of Bayley from the photoboard, she said, “It definitely looks like No. 4” (Bayley v The Queen, 2016, at [41]). By the time of this photoboard identification, GH had seen photos of Bayley multiple times, and she knew that he had been
charged with the rape and murder of Gillian Meagher, as well as multiple charges against GH herself, including false imprisonment, assault, and rape.

On the 14 July 2015, Adrian Bayley was convicted by the County Court of Victoria of the offences against GH. In the original trial, the defence counsel sought to exclude the identification evidence on the basis of section 137 of the Evidence Act 2008 (Vic): “evidence should not be admitted if its probative value is outweighed by the danger of unfair prejudice to the defendant” (s 137). The judge refused to exclude the evidence, ruling that the probative value of the Facebook identification was not outweighed by unfair prejudice, and that the photoboard identification had its own probative value. The County Court of Victoria’s approach to section 137 aligns with Dupas v The Queen (2012)—the jury can accept evidence as credible without accepting it as reliable. In other words, the jury can hear the identification evidence but choose to not convict based on that evidence.

The Court of Appeal Judgement

On appeal against the DPP v Bayley (2015) decision, the defence argued that the conviction was based on unsafe or unsatisfactory grounds; that is, that the evidence was too problematic and the jury could not have reasonably rendered a guilty verdict based on such evidence. The Victorian Court of Appeal held that the identification evidence should not have been admitted in the first place and took a different approach to section 137, consistent with the High Court decision in IMM v The Queen (2016). It determined that while a trial judge should proceed on the assumption that the jury will accept the evidence, and should not have regard to the credibility or reliability of the evidence, the trial judge should nevertheless exclude the evidence if it is not merely weak but “simply unconvincing” – it its probative value is scant.
In their decision, the Court of Appeal detailed six problems with the identification evidence:

1. The initial identification was made from a single photograph seen on Facebook.
2. The photograph from which the Facebook identification was made was taken 11 years after the rape.
3. The identification was made 12 years after the rape.
4. The circumstances of the identification were suggestive because GH was aware that Adrian Bayley was under arrest for the rape and murder of Gillian Meagher.
5. The victim may have selected the photograph of Adrian Bayley from the formal police photoboard as a result of the displacement effect. In other words, GH’s memory for the original perpetrator may have been contaminated after seeing Bayley’s image on Facebook, which may have influenced her identification during the photoboard.
6. The photoboard identification provided evidence of similarity—not identification—because when she selected Bayley, GH said, “It [the perpetrator] definitely looks like number four [the photo of Adrian Bayley]” (Bayley v The Queen, 2016, at [58]).

After weighing the probative value of the identification evidence, the judge concluded that it should have been excluded in the initial trial. Because there was no other evidence, the Court of Appeal acquitted Bayley on all charges in GH’s case.

**Issues with the Identification Evidence**

In their decision, the Court identified several important factors that can undermine the reliability of identification evidence. In this article, we will turn to psychological literature to
seek answers to issues outlined by the Court of Appeal. What does the empirical research tell us about the influence of delay on identification accuracy? How reliable are identifications made from a single photograph? Does knowledge of the suspect’s background influence the decision-making process? How does the displacement effect (seeing someone’s face before the identification procedure) affect the reliability of an identification? In addition to briefly reviewing the psychology of eyewitness identifications, we will discuss factors that may have influenced the jury’s views as to this weak evidence, leading them to render a guilty verdict in the initial trial. We would like to emphasise that the aim of this article is to present the psychological research relevant to this case, not to provide a legal analysis of the court’s decision. Critically, we are not questioning GH’s report of the assault or her decision to not initially report it, nor are we making any judgement about Bayley’s involvement in her assault; rather, we present the empirical evidence to explain why the identification evidence was not strong enough to uphold Bayley’s conviction.

In an ideal world, memory would work like a recording device, copying events, storing them without error, and reproducing them without modification. The reality is, however, that memory is a reconstructive process; memories are not preserved in an unchanged way. Memory is not a simple process. It includes three stages: encoding, storage, and retrieval and at each stage, information can be lost or contaminated (Foster, 2009). We present the issues with the identification evidence in this case in line with this conceptualisation of memory. Differentiating the three stages of memory is important because it allows us to see how GH’s memory may have been affected by outside factors—both before the police were contacted (e.g., the length of the delay or circumstances of the Facebook identification) and once the police became involved (e.g., how the photoboard identification procedure was conducted). This distinction between
factors under and outside of the control of the criminal justice system is crucial because the criminal justice system can influence only those factors under their direct control (Wells, 1978).

**Issues at the Encoding Stage**

The first stage, encoding, occurs when a person is first exposed to the material (Foster, 2009). Several factors can influence the quality of memory at the encoding phase, including: how much attention the witness paid to the perpetrator’s face; how long the witness saw the perpetrator (a.k.a., exposure duration); the quality of witnessing conditions (such as lighting and distance); and amount of stress experienced by the witness. Twelve years after the attack, it is difficult to reconstruct the quality of witnessing conditions (e.g., the lighting, the exposure duration, etc.). What is certain, however, is that experiencing a violent sexual attack is undoubtedly an extremely stressful experience. During the attack GH testified that she was “pleading [for her] life” and “was horrified and in agony” (*Bayley v The Queen*, 2016, at [30]).

**Stress.** Although people are unlikely to forget that a stressful event occurred, the questions of practical importance are: How accurate are people’s memories about the central details of that stressful event? Do people memorise the face of a once-seen perpetrator? In response to these questions, the research indicates that memories for highly-stressful events are generally poor. A meta-analysis of 27 tests of eyewitnesses’ identification performance indicated that stress levels are negatively correlated to identification accuracy (Deffenbacher, Borstein, Penrod, & McGorty, 2004). Eyewitnesses who experience high levels of stress make more false identifications compared to low stress circumstances. Although the evidence that stress impairs identification accuracy is robust, these results are based on experiments conducted in the laboratory under controlled conditions. One could argue that these laboratory results may not be indicative of the workings of memory in real-life.
One of the most realistic studies to date with unequivocally stressful conditions supports these laboratory findings. Five hundred military personnel were interrogated during a survival-training program designed to imitate the prisoner-of-war experience. All participants were interrogated under high- and low-stress conditions. In the high-stress condition, an interrogator physically confronted the participants; in the low-stress condition, another person interrogated them without physical confrontation. The interrogations lasted a total of 40 minutes each, with 4 hours between them. When the interrogation did not involve physical confrontation (i.e., low-stress condition), 76% of participants correctly identified the interrogator from a simultaneous photoboard; in contrast, only 34% made an accurate identification when the interrogation was highly stressful. Similarly, 68% of participants in the high stress condition made a false positive identification, compared to only 12% in low stress condition (Morgan et al., 2004). Similar results were obtained in another study that tested the accuracy of eyewitness identifications under stress. Valentine and Mesout (2009) recorded participants’ heart rate and levels of anxiety during their visit to the London Dungeon. During the visit, participants encountered a “scary person” who they were later asked to describe and identify from a 9-person simultaneous target-present photoboard. Out of all participants who experienced high levels of state anxiety, only 17% correctly identified the “scary person”, compared to 75% of correct identifications among those who had lower state anxiety scores.

Memories for highly stressful events are also more prone to incorporate false information. When participants in the survival-training course were presented with the photo of another person between the time of the interrogation and the presentation of the photoboard, 91% subsequently mistakenly identified that innocent person as their interrogator, even though the
actual interrogator was not present in the photoboard (Morgan, Southwick, Steffian, Hazlett, & Loftus, 2013).

**Issues at the Storage Stage**

The storage stage is the period between the initial exposure and later recollection of that information. Events that take place between the crime and identification procedure (during the storage stage) can also influence an eyewitness’s memory, because memory, once encoded, is still susceptible to extraneous influences. Memories can fade if there is a long delay between the crime and the identification, but they also can be replaced or altered if an eyewitness discusses the event with other witnesses, police officers, or reads media reports.

**Delay.** GH identified Adrian Bayley 12 years after her attack. Common sense suggests that long delays may decrease identification accuracy, but what does the research tell us? To answer this question, we turn to studies examining people’s memory for once-seen faces after a delay. Most eyewitness studies investigating delay have varied the time between the exposure to the perpetrator and the identification attempt (usually from a photoboard) from one day to one week. A meta-analysis of 53 studies found that long delays reduced people’s ability to identify once-seen faces (Deffenbacher, Bornstein, McGorty, & Penrod, 2008). Shepherd (1983) reported the longest delay tested to date: 11 months. After a one-week delay, 65% correctly identified the perpetrator; in contrast, after an 11-month delay, only 10% made an accurate identification.

The majority of these studies, however, focused on identifications made from photoboard, a procedure that differs from the conditions of GH’s Facebook identification from a single photo. Research demonstrated that accuracy of identifications from a single photo (i.e., a showup) decreased within 24 hours (Dysart & Lindsay, 2007). Although no research examined the effects of a 12-year delay on a person’s ability to recognise a perpetrator from Facebook,
there is little reason to suspect that an identification made under these conditions would produce higher accuracy rates, given that memory consistently degrades over time. One could reasonably conclude that—12 years after the crime—GH’s memory for the perpetrator would certainly have been less reliable than it would have been immediately after the offence.

**The influence of expectations.** When GH identified Bayley from Facebook, she was already aware of Gillian Meagher’s death and knew that Bayley had been arrested for her rape and murder (*Bayley v The Queen*, 2016). Knowledge of a suspect’s background can influence the expectations held by the witness and skew the decision-making process, increasing the probability of a false identification. Confirmation bias refers to the tendency to seek information that confirms one’s expectations and beliefs while disregarding all other information (Plous, 1993). This bias occurs automatically, emerging regardless of a person’s wishes or attempts to be impartial (e.g., Nickerson, 1998; Findley & Scott, 2006; Kunda, 1990). This bias can be observed in a variety of tasks, but for the sake of brevity, we will focus on the ones most relevant to the forensic context.

When making identification decisions, eyewitnesses may be influenced by various factors, rather than basing their decisions solely on simple recognition. Factors such as beliefs about the case or hints given by the police officer during the identification procedure can skew an eyewitness’s decision process. In the field-simulation study examining eyewitnesses identifications from showups, false identification rates were higher when eyewitnesses believed that they witnessed an actual crime (40%), compared to the laboratory conditions (13%; Eisen, Smith, Olaguez, & Skerritt-Perta, 2017). This result highlights the influence that emotions and expectations can have of decision making.
In one study, participant-eyewitnesses were exposed to a crime video and asked to make an identification from a photoboard that did not contain the perpetrator from the video. After their identification decision, some participants were told that another photoboard member (not the one they chose) confessed. The majority (61%) of participants changed their decision and instead identified the alleged confessor. Moreover, half of the participants who did not initially make an identification abandoned their original correct rejection and chose the alleged confessor (Hasel & Kassin, 2009). This finding highlights the lack of independence of evidence and illustrates that eyewitnesses may be willing to identify a person even if they do not recognise him. This situation is dangerous because eyewitnesses may become invested in a choice that can replace their original memory altogether (see the section on the displacement effect).

Expectations can also influence eyewitnesses’ behaviour in more subtle situations. For example, when eyewitnesses receive pre-identification instructions suggesting that the perpetrator is in the photoboard, they are more willing to choose someone, compared to when they are warned that the perpetrator may or may not be present (for a meta-analysis of 18 studies see Steblay, 1997; for a review see Steblay, 2013). This willingness to choose—and the associated consequences—is exacerbated when the photoboard does not contain the perpetrator. Furthermore, the photoboard administrators’ expectations can also affect eyewitnesses’ choices. When the administrator knows the suspect’s identity, eyewitnesses are more than twice as likely to choose the suspect than any other photoboard member, regardless of the actual guilt of the suspect (Austin, Zimmerman, Rhead, & Kovera, 2013; Canter, Hammond, & Youngs, 2013) and they are more confident in that decision (Charman & Quiroz, 2016). In a broader context, the influence of expectations on decision-making has also been studied in forensic science (for a review see National Research Council, 2009). If the forensic expert is aware that the suspect
confessed, the probability of finding a match between a sample found at the crime scene and a sample from the suspect increased. This effect was observed in a handwriting comparison task (Kukucka & Kassin, 2012), a fingerprint comparison (e.g., Dror, Peron, Hind, & Charlton, 2005), and even DNA testing (Dror & Hampikian, 2011).

In the present case, the details of the Facebook identification are unknown. It is unclear whether GH was scrolling that page to identify her attacker or if she was looking at that page for the details of Gillian Meagher’s case or for some other purpose. What is certain, however, is that—when she first saw his photo—GH knew of Bayley’s involvement in Gillian Meagher’s case. Thus, this knowledge may have influenced her expectations of Bayley and predisposed her to identify him as the person who raped her.

**Issues at the Retrieval Stage**

Retrieval is a process of actively searching for information. Every time a person tries to remember something, he or she is actually rebuilding or reconstructing that memory. It is during this reconstruction that errors can occur because information from other sources can be incorporated into memory. Some of the factors that can affect accuracy at this stage are how the identification procedure was conducted or what instructions the officer gave to the witness prior to conducting a photoboard. These factors can affect accuracy at future retrievals, as well as influencing the witness’ confidence in that memory. In the present case, these factors include: the initial Facebook identification was made from a single photograph, which could influence an eyewitness’s willingness to identify a person shown in the photo as the perpetrator; the memory for the perpetrator could have been displaced by Adrian Bayley’s face; the photo was taken 11 years after the attack; and the appearance of the perpetrator might have changed.
**Single photograph identification.** GH initially identified Bayley from a single photograph on Facebook, which is similar to a showup procedure. In a show-up, the police present the eyewitness with a single suspect or a photo of the suspect and ask the eyewitness to decide whether this person is the perpetrator (Dysart & Lindsay, 2007). Why is a show-up problematic? First, it immediately reveals to the eyewitness the identity of the suspect and influences the eyewitness’s expectations that the suspect is likely to be the perpetrator (Dysart & Lindsay, 2007). Second, if presented with the show-up, the eyewitness has only two options—identify the person or reject the photo. Thus, a show-up produces more identification errors compared to a photoboard (Gronlund et al., 2012). Moreover, even immediate show-ups still perform worse than delayed photoboards (Wetmore et al., 2015). Show-ups are particularly dangerous for innocent suspects, with eyewitnesses making more false identifications from show-ups compared to photoboards (Steblay, Dysart, Fulero, & Lindsay, 2003). This risk increases with repeated showups (Smith et al., 2014).

In contrast, in a photoboard an eyewitness is shown the photos of a suspect and several foils (known-innocent people) and is unaware of which person is the actual suspect. The addition of foils serves two important functions. First, it reduces the suggestiveness of the procedure because an eyewitness does not know which person is the police’s suspect. Second, if an eyewitness chooses a foil, it demonstrates their willingness to choose from the photoboard even in the absence of a strong recognition (Gronlund et al., 2012). In Bayley’s case, however, the photoboard did not serve any of these functions because GH was aware of the identity of the suspect and had previously identified him from Facebook.

**Displacement effect.** The Court of Appeal argued that a displacement effect could have taken place, in that GH’s initial memory of the perpetrator may have been displaced with an
image of Adrian Bayley. The displacement effect was originally described in *Alexander v The Queen* (1981):

> Having been shown a photograph, the memory of it may be more clearly retained than the memory of the original sighting of the offender and may, accordingly, displace that original memory. Any subsequent face-to-face identification, in court or in an identification parade, may, on the identifying witness’s part, in truth involve a matching of the man so identified with the remembered photograph, which has displaced in his memory his recollection of the original sighting. (p. 409)

Although common in legal rulings, the term “displacement effect” is not used in psychological literature. Instead, psychologists would categorise the “displacement effect” as the *mugshot exposure effect*. A mugshot exposure (i.e., viewing a photograph of a person prior to the identification procedure) decreases eyewitnesses identification accuracy (for a meta-analyses of mugshot exposure effects, see Deffenbacher, Bornstein, & Penrod, 2006). However, the term that best describes a situation when an eyewitness makes repeated identification decisions is the *mugshot commitment effect*. When an eyewitness identifies someone, they are more likely to choose the same person again, even if this person is not the actual perpetrator. An eyewitness chooses the person again because they are committed to their initial identification decision (Dysart, Lindsay, Hammond, & Dupuis, 2001). Although the consequences of the mugshot exposure and the mugshot commitment effects are similar (i.e., the wrong person has been selected as the perpetrator), the mechanisms driving the eyewitness’s decision are different. In the case of mugshot exposure, the memory of one face replaces the memory for the perpetrator’s face. In the case of commitment, however, the eyewitness is making the erroneous decision because they had previously identified the same person.
The commitment effect demonstrates that repeated identification decisions are not independent: an earlier decision can influence subsequent ones. This means that one identification procedure cannot be used to confirm the eyewitness’s identification from a previous procedure. Any correspondence between decisions from the two procedures may simply be the result of the commitment effect, and should not be interpreted as evidence of the eyewitness’s accuracy (Lawson & Dysart, 2015). In the Bayley case, GH made an initial identification decision before the formal photoboard, potentially resulting in the commitment effect.

**The photo was taken 11 years after the attack.** The Court of Appeal noted that the age of Adrian Bayley’s photo at the time of the Facebook identification is one of the factors that diminished the probative value of the evidence. This is an important consideration because a change in appearance (e.g., disguise, ageing) can affect identification accuracy (e.g., Laughery, Alexander, & Lane, 1971; Mansour et al., 2012; Wells & Hryciw, 1984). As one might expect, similarity between the perpetrator’s appearance at the time of exposure and at the time of identification increases the probability of making a correct identification (Molinaro, Arndorfer, & Charman, 2013). Changing the style of the perpetrator’s facial hair or hairstyle can impair identification accuracy (Patterson & Baddeley, 1977; Patterson, 1978; Pozzulo & Marcinak, 2007). One could argue, however, that Adrian Bayley’s appearance did not significantly change between 2000 and 2011 (see figure 1); thus, this factor may have had little influence on GH’s decision to identify him.
Evidence of Similarity. When GH identified Adrian Bayley from the photoboard, she said, “It [the perpetrator] definitely looks like number four [the photo of Adrian Bayley]” (Bayley v The Queen, 2016, at [41]). The Court of Appeal used this statement to influence their interpretation of the weight of this evidence; they deemed her selection to be evidence of similarity rather than evidence of a positive identification. The Australian Law Reform Commission Report (2006) describes evidence of identification as, “Where a witness claims to recognise the defendant as the person seen on the relevant occasion” (section 13.11). Whereas evidence of similarity (or evidence of resemblance) is described as, “Evidence that a person shares certain features or attributes in common with the accused or that he or she looks like the accused” (section 13.11). This raises the question of whether the choice of words with which an eyewitness makes an identification is indicative of the weight of the evidence. To date, no eyewitness studies, to our knowledge, have directly investigated whether the words that the eyewitness uses when making an identification can be indicative of one of two types of evidence: identification vs. similarity.
evidence. Nonetheless, a relevant line of research to investigates the decision processes eyewitnesses use to reach their decisions, and how these processes relate to identification accuracy.

Wells (1984) suggested that when an eyewitness is presented with a photoboard, he or she can engage in one of two decision strategies: relative or absolute judgments. The relative judgment strategy describes the process in which an eyewitness compares several photoboard members with each other, trying to find the closest match to their memory among those present. An eyewitness who uses relative judgment chooses the photoboard member who looks most like the perpetrator among other members, provided he is a close enough match to the witness’ memory. This is a problematic strategy if the photoboard does not, in fact, contain the actual perpetrator. Thus, it puts an innocent suspect at risk if he happens to look more similar to the perpetrator than the other photoboard members (Wells, 1984; Wells, 1993). An absolute judgment, on the other hand, refers to a process in which an eyewitness identifies the photoboard member only when this member is a good match to the eyewitness’s memory; thus, there are no between-member comparisons. Wells (1984) argued that an absolute judgement reduces the risk of false identification, because an eyewitness makes a selection only when the suspect is a good match to his or her memory of the perpetrator, not when the suspect is simply the best match among the presented alternatives.

Other researchers have conceptualised the eyewitness identification decision process as being automatic or deliberative (Dunning & Stern, 1994). They measured underlying decision processes by asking eyewitnesses how they reached their decision. Endorsing statements reflecting comparisons between line-up members or engaging in a process of elimination suggests a more deliberative strategy rather than automatic recognition (Dunning & Stern, 1994).
Accurate eyewitnesses tended to report automatic recognition of the perpetrator (i.e., that the person ‘popped out’ to them), whereas inaccurate eyewitnesses tended to engage in a deliberate process of eliminating other options, until they found a good enough match.

This distinction between automatic recognition and process of elimination resembles the Courts’ distinction between evidence of identification and evidence of similarity, respectively. However, to date there is no empirical evidence that could show the distinction between evidence of similarity and evidence of identification. In the present case, when GH chose the photo of Adrian Bayley, she said, “it definitely looks like...,” but did not indicate that she based this decision on a comparison amongst photoboard members (e.g., “it looks most like...”). As such, it is difficult to evaluate the Court of Appeal’s reasoning in light of the current scientific evidence. More research is necessary to support the Court’s willingness to give more weight to evidence of identification than evidence of similarity by exploring how the two conceptualisations relate to identification accuracy.

**Summary and Additional Considerations**

As we mentioned, it is important to differentiate between factors that were not under the control of the criminal justice system (e.g., the length of the delay or circumstances of the Facebook identification) and those that can be controlled (e.g., how the photoboard identification procedure was conducted). Nonetheless, evidence obtained under suggestive circumstances can influence jurors’ judgments and their willingness to convict the defendant.

The present case illustrates that memory is fragile and at each of the three stages of memory multiple factors can undermine the accuracy of memory and, consequently, the reliability of the identification evidence. We agree with the Court’s ruling that GH’s photoboard identification should not have been admitted as evidence against Bayley because the circumstances under
which the identification was obtained were overly suggestive. We now turn to another matter—
namely the jury’s decision to convict Bayley at the initial trial despite the weak evidence against
the defendant.

**Why Did the Jury Render a Guilty Verdict?**

An impartial jury ensures that the defendant’s right to a fair trial is met. The notion of the
jury’s impartiality—basing their verdict on the evidence presented at trial, not on beliefs or
information obtained outside the courtroom—is one of the basic assumptions of the Australian
criminal justice system. However, the question remains, can jurors ignore extra-legal information
and base their decisions only on evidence presented at trial? The jury convicted Bayley despite
multiple weaknesses in the identification evidence and defence’s attempt to educate jurors by
asking an expert psychologist—Associate Professor Richard Kemp—to testify about the
problems with the identification evidence.

The Court of Appeal disagreed with the jury’s verdict and argued that the evidence was
too weak to support a guilty verdict. *Bayley v The Queen* (2016) ruled that, “Having made our
own independent assessment of the evidence as a whole, we are firmly of the view that a
properly instructed jury, applying themselves appropriately to their task, must have had a
reasonable doubt as to the applicant’s guilt. GH’s identification evidence was so weak that no
other conclusion was reasonably open” (at [99]).

Eyewitness evidence is extremely persuasive to jurors (Boyce, Beaudry, & Lindsay, 2007). The mere presence of such evidence—regardless of its reliability—can drastically
increase guilt ratings (Loftus, 1974). People attribute disproportionate weight to eyewitness
evidence partially because they overestimate the accuracy of eyewitness memory. For example,
Brigham and Bothwell (1983) presented descriptions of previously-conducted studies and asked
participants to predict the accuracy of those eyewitnesses. The vast majority of participants overestimated the eyewitnesses’ accuracy rates, illustrating “overbelief” of eyewitnesses and a general lack of knowledge of factors that influence eyewitness accuracy. Even when jurors were aware of the suggestiveness of an identification procedure, they still believed the eyewitness evidence (e.g., Devenport, Stinson, Cutler, & Kravitz, 2002). Thus, although people may be able to distinguish between fair and suggestive procedures, they do not use this information when rendering a verdict. The persuasiveness of eyewitness evidence, coupled with a lack of knowledge of factors that influence eyewitness evidence and a disregard of the dangers of suggestive procedures, increases the risk of juries basing their decision on unreliable evidence.

Furthermore, by the time of GH’s case, Adrian Bayley had already been convicted of the rape and murder of Gillian Meagher, whose tragic death attracted significant media attention. Thousands of people participated in a march in memory of Gillian Meagher on the 30th of September, 2013 (Zielinski, 2013). Even though GH’s case was heard in 2014, it is highly probable that jurors still remembered the circumstances of Gillian Meagher’s death and that Adrian Bayley pleaded guilty to her murder. This knowledge might have influenced jurors’ expectations and skewed their perception of the defendant in favour of the prosecution before the trial had even begun.

Pre-trial publicity exposes jurors to information about the case that may not be true or may not be admissible in court, such as non-existent evidence or the defendant’s history of prior offences. This pre-trial information may irreversibly bias jurors’ perceptions of the case. Subsequent information presented within the court case itself may be perceived through the lens of the pre-trial publicity. Research into the effects of pre-trial publicity found that it increased jurors’ willingness to convict the defendant (for a meta-analysis, see Steblay, Besirevic, Fulero,
& Jimenez-Lorente, 1999), decreased the perceived credibility of the defendant, and elicited a negative emotional response from jurors (Ruva, Guenther, & Yarbrough, 2011). This effect was even stronger in cases involving sexual assault and murder than in other types of cases (Steblay et al., 1999). The influence of pre-trial publicity is so strong that not even traditional legal safeguards, such as jury selection or judicial instructions, have been unable to overcome its biasing effects (Kramer et al., 1990; Lieberman & Arndt, 2000). Moreover, the deliberation process that is believed to serve as a safeguard against any individually-held prejudices did the opposite by increasing the damaging effects of pre-trial publicity and leading to more guilty verdicts than reported in pre-deliberation votes (Kramer et al., 1990). Furthermore, social media (e.g., Facebook) and the internet can also compromise the independence of jury decision making as jurors can search for information about the case despite being instructed not to do so.

Identification made by GH was persuasive to the jury, because she testified about her firsthand experience and jurors can be insensitive to factors that undermine the reliability of an identification. Moreover, given the widespread media attention received by this case, it is possible that the jury still remembered who Adrian Bayley was and that knowledge affected their decision-making. Examples such as this illustrate a need for revision of the Criminal Justice System’s perceptions of jurors as purely rational decision makers and a need to focus on ways to improve accuracy of jurors’ evaluation of eyewitness evidence.

Discussion

Our review highlighted several issues regarding the probative value of the identification evidence in Adrian Bayley’s trial. For the most part, the psychological research supports the issues identified by the Court of Appeal (i.e., delay, change of appearance, single photograph identification, and the influence of expectations). However, two issues required additional
First, the Court raised concerns about the possibility of GH’s memory for the perpetrator being displaced by Bayley’s photo during the Facebook identification. In the legal literature it is known as the displacement effect. However, in psychological literature a similar phenomenon is called a mugshot exposure. The difference in terms that appear to be describing the same effect may contribute to confusion and miscommunication between legal practitioners and psychological scholars. Moreover, in the case of GH’s photoboard identification, the mugshot commitment effect offers the best explanation. GH did not simply see Adrian Bayley, she identified him from the Facebook photograph, a situation that is different from the traditional mugshot exposure effect. Second, the distinction between evidence of identification and evidence of similarity based on the choice of words with which an eyewitness makes an identification lacks empirical support. Although an eyewitness’s self-reported decision strategy may correlate with their identification accuracy there is conflicting research about its utility (e.g., Ross, Benton, McDonnell, Metzger, & Silver, 2007) and it is not yet clear how these decisions processes correspond to the distinction between two types of evidence.

A separate issue was the jury’s decision to convict Bayley at the initial trial based on the weak identification evidence alone. Lay people may not possess appropriate knowledge about factors that affect eyewitness accuracy and are often convinced by this type of evidence regardless of the quality of witnessing and identification conditions. The media coverage surrounding Bayley may have also played a role in this case. Jurors may find it difficult to make decisions independent of information they saw or heard outside of the courtroom. Even when instructed not to do so, they still have a tendency to rely on information heard elsewhere when deciding the case. Future research must find effective ways of educating jurors and/or increasing the role of judges as gatekeepers for this type of evidence.
Bayley v The Queen (2016) adds to a number of other cases in which a Facebook identification was an issue. Identifications made from Facebook provide new challenges for the criminal justice system. The decisions are made outside of the police station, without the benefit of any of the traditional safeguards used to ensure the fairness of the procedure. Moreover, an initial Facebook identification may contaminate an eyewitness’s memory for the perpetrator and significantly increase the risk of the commitment effect in a subsequent formal identification procedure. How are these emerging Facebook identifications currently dealt with in Australia? McGorrery (2015) provided an overview of five recent cases that discussed such identifications; only one case rendered Facebook identifications inadmissible (Strauss v The Police, 2013). In four other cases, the Facebook identification evidence was admitted, even though these identifications were made under highly suggestive circumstances in the absence of any safeguards.

These leading cases are likely to influence how the police and courts approach Facebook identifications in future cases. In his recommendations, McGorrery (2015) underlined the importance of collecting initial statements from eyewitnesses at the crime scene and warning them not to use Facebook. Obtaining an uncontaminated confidence statement is of utmost importance, as the confidence at the time of the identification can be indicative of an eyewitness’s accuracy (Brewer & Wells, 2006; Wixted & Wells, 2017). It is unclear, however, whether the confidence-accuracy relationship has similar utility in cases of Facebook identification. McGorrery (2015) also argued that the formal identification procedure should be conducted, even if an eyewitness identified the suspect from Facebook. He noted that even though there is a high risk of the displacement effect, the formal procedure can highlight the unreliability of the eyewitness if he or she fails to identify the suspect from the photoboard.
(Strauss v The Police, 2013). In our opinion, however, any subsequent procedure should not be admitted in court, because it may be the result of the commitment effect rather than an indication of a strong memory for the suspect as the perpetrator (see Steblay & Dysart, 2016).

**Conclusion**

This review demonstrates that lawyers and psychologists may reach the same conclusions, but use different reasoning and evidence to inform their decisions. Judges base their decisions on the events observed in the case before them and use legal precedents. Psychologists, on the other hand, rely on findings from experimental studies that are generally accepted by the field. Some of the Court of Appeal’s case observations correspond to the findings of psychological studies, but some do not. This case highlighted some of the issues that are not yet addressed by psychological science. It is unclear how Facebook identifications differ from the traditional form of identifications and what is the best way to communicate with eyewitnesses in such cases. The difference between evidence of similarity and evidence of identification is another gap that psychologists should fill. We would also like to underline the need for continued interaction between psychological researchers and legal practitioners. It is important to establish ongoing communication to ensure the consistency and effectiveness of legal decisions and the applicability and relevance of psychological research.
Acknowledgments

This work was partially supported by the Swinburne Postgraduate Research Award to the first author. We thank Professor James Ogloff for his comments on the earlier version of this article and for his guidance and expertise. We would like to thank Dr. Mehera San Roque for her advice. We are grateful to Professor Mirko Bagaric for his suggestions. We also thank Dr. Michelle Noon and Mr. Jamie Smith-Morvell for their thorough reviews and insightful suggestions.

Declaration of Interest

Alena Skalon has declared no conflicts of interest. Jennifer L. Beaudry has declared no conflicts of interest.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.
References

Alexander v The Queen (1981) 145 CLR 395

ALRC Report 102 (2006, February, 8). Retrieved from

administration. In B. L. Cutler (Ed.), Reform of eyewitness identification procedures (pp.

Bayley v The Queen [2016] VSCA 160

evidence. In R. Lindsay, D. Ross, D. Read, & M. Toglia (Eds.), Handbook of Eyewitness
Psychology: Memory for People (pp. 501–525). Mahwah, NJ: Lawrence Erlbaum and
Associates.

identification: effects of lineup instructions, foil similarity, and target-absent base rates.


Charman, S. D., & Quiroz, V. (2016). Blind sequential lineup administration reduces both false
identifications and confidence in those false identifications. Law and Human Behavior,


*DPP v Bayley* [2015] VSCA 698, conviction trial 1


http://dx.doi.org/10.1016/j.scijus.2011.08.004


Dupas v The Queen (2012) 40 VR 182


[http://dx.doi.org/10.1037/0033-2909.108.3.480](http://dx.doi.org/10.1037/0033-2909.108.3.480)


Strauss v. The Police (2013) 115 SASR 90


271

http://dx.doi.org/10.1016/j.jarmac.2015.11.001


Appendix 2. Victorian Judicial Instructions

Identification is an important issue in this case. The case against the defendant depends, to a significant extent, on evidence claiming to identify the defendant from a photoboard.

Before you decide whether to accept this evidence, I must give you some warnings about identification evidence.

Before I give you these warnings, and to help you understand the reason why I am giving you these warnings, I will briefly explain how people are identified.

There are three stages that are involved whenever a positive identification is made and a witness may make an error at each stage. First, the witness must have observed somebody during the crime. Second, the witness must have retained an image of that person in his or her mind until the time of the identification. Third, the witness must have later seen the defendant, or a picture of the defendant, and identified him or her as being the person seen who committed the crime.

Identification evidence is potentially unreliable. For that reason, you must exercise caution in determining whether to accept the evidence and, if you do accept it, the weight that you accord to that evidence.

One of the reasons that identification evidence is potentially unreliable is that while a witness may honestly believe that his or her evidence is accurate when he or she is actually mistaken. And the mistaken evidence of a witness may be convincing.

The experience of the law has shown that witnesses have given mistaken identification evidence which has resulted in innocent people being convicted.

I will now turn to discuss the significant matters in this case which may make the identification evidence unreliable.

Having given you that general warning about identification evidence, I now want to look at some of the specific factors that may affect the reliability of the evidence in this case.

Here, there are several factors that may be relevant to your assessment of the reliability of the identification evidence. These are the circumstances in which the offender was observed; the characteristics of the witness who gave evidence; and the way in which the accused was identified.

You should examine each of these factors closely when deciding whether to accept the identification evidence.

I will now look at these factors in more detail.

Circumstances of Observation

You should examine the circumstances in which the offender was observed. You should consider what opportunity for accurate observation existed. Some of the questions you should ask yourself include:

For how long did the witness observe the person?
How far away was the witness from what s/he was observing?
What was the angle of observation? For example, did the witness see the person’s face or only his/her back?
Had the witness ever seen the person s/he was observing before?
What was the light like?
Did anything get in the way of the witness’s view, such as passing people or traffic?
Did the witness have a reason for trying to observe the person involved and to remember his/her characteristics?
Did the person observed have any distinguishing features or characteristics which would make it likely that the witness would remember him/her? For example, did s/he have a scar or a tattoo?

**Factors Concerning the Witness**

You must also consider the characteristics of the witness who gave the identification evidence. In that context, some of the questions you should ask yourself include:

- Is it possible to assess the quality of this witness as an observer?
- Was the witness stressed or fearful at the time of the observation? If so, what effect would this stress have had on him/her? For some people, their powers of observation increase under stress. Others "black out" and their powers of observation diminish. You need to decide how the witness is likely to have reacted in this case.
- Were there any other factors that could have affected the witness’s powers of observation, such as drugs or alcohol consumption, or fatigue?

**Factors Concerning the Identification**

You must consider the way in which the defendant was identified/similarities between the person observed and the defendant were noted. Some of the questions you should ask yourself include:

- Did the witness give a description of the offender before identifying the accused? If so, does the description match the accused?
- Is the witness relying too heavily on a particular memorable feature or characteristic of the accused in identifying him/her?
- How long was there between the incident and the identification? Was it likely that the witness’s memory was affected by any delay?
- Was the identification process conducted fairly? For example, did the other people in the photoboard look sufficiently similar to the accused?
- Did the witness hear a description or see a picture of the accused before attempting to identify the offender?
- Was the witness influenced in any other way to identify the accused – for example, by the behaviour of the police?

In this case, as the defendant was identified from a photograph, I need to give you an additional warning about photographic identification evidence.

This sort of evidence may be unreliable due to the differences between photographs and real life. For example, photographs are two-dimensional, and do not show the way a person moves, the range of their facial expressions, their body shape, or many of the other characteristics that can help you to identify a person.

The photograph used to identify the accused may also have been taken in very different circumstances from those in which the offender was observed. For example, the light in the photograph may be much better than it was at the time of the crime.

These factors can increase the risks of misidentifying the offender, who may look like the accused as seen in a photograph, but may look different when viewed face-to-face. You should therefore treat photographic identification evidence with special care.

There is an additional problem with photographic identification. As the accused was not present during the identification process, he is unlikely to have any first-hand information about the way in which his/her photograph was selected. Instead, s/he can only rely on the cross-examination of the people who were present to gain any information about the conditions in
which the identification took place, and what safeguards against error were taken. You should take this disadvantage into account when considering the evidence.

You may have noticed that the defendant was identified from a photograph held by the police. You are not to attach any significance to this fact. The police have photographs of many different people for a variety of reasons. You must not assume that, because the police had a photograph of the defendant, s/he has a criminal record or has previously been charged with an offence. In fact, you must not draw any conclusions from the fact that the police had a photograph of the defendant.

**If an Identification Parade Could Reasonably Have Been Held**

In this case, it would have been possible for the police to hold an identification parade instead of having the eyewitness identify the defendant from a photograph.

In an identification parade, a witness is asked whether they can identify the offender from a selection of people resembling the accused. This process has two main advantages over identification from photographs.

First, the witness is identifying an actual person, rather than a two-dimensional representation of that person. S/he is able to see all of the accused’s physical characteristics, such as the way s/he moves and his/her facial expressions. This makes it more likely that the witness will accurately identify the offender.

Secondly, as the accused is present at an identification parade, s/he is able to obtain first-hand information about how it is carried out. S/he will be able to see what steps are taken to make sure that it is conducted fairly, rather than having to rely on cross-examination of the people present.

In this case, the defendant was deprived of these benefits. He may also have lost any advantage s/he might have gained if an identification parade was inconclusive. You should take these disadvantages into account when assessing the evidence against the defendant.

Finally, you should consider any other significant factors that may affect the reliability of the identification evidence.

**Summary**

To summarise, it is important that you take care in determining whether you accept identification evidence, and if you do accept it, in deciding what weight to give to that evidence.

If, after careful examination of the identification evidence, and in light of all of the circumstances and other evidence given in the case, you find that the accused was correctly identified, then you can use the evidence in reaching your verdict.
Appendix 3. Full Trial Transcript used in Paper 2

**R v. Roger A. Harvey**

**JUDGE’S CASE SUMMARY**

The following transcript concerns the case of Roger A. Harvey (a 32 year-old male carpenter) who has been charged with the armed robbery of the store clerk. The witness was at a Seven-Eleven store in Melbourne, Victoria at about 10:30 p.m. on October 23, 2015, when a man entered the store. A man pulled out a knife and ordered the store clerk to open the register. The store clerk then filled a black rubbish bag with cigarettes. When the robber left, the store clerk called 000. The witness saw a white ute pull out of the parking lot. Police officers responded to the call within minutes. They asked the witness to describe the robber and the vehicle he saw. The store clerk also reported that the robber took $220 from the register and 30 packs of cigarettes.

A few kilometers from the store the police found a white ute parked on the side of the road. In it they found 8 packs of cigarettes of various brands, a black rubbish bag, and a blue jumper. The ute was registered to the defendant, who matched witnesses' description of the robber. Two days later the witness saw a photoboard and identified the defendant, Roger Harvey, as the robber.

**Mr. Walker (prosecutor):** The victim was working at a Seven-Eleven store when a robber arrived threatening him with a knife. The robber ordered him to fill a black rubbish bag with cigarettes and to open the register. The defendant’s white ute was found near the crime scene. The ute contained 8 packs of cigarettes of various brands, a black rubbish bag, and a blue jumper. The ute was registered to the defendant, who matched description of the robber. The witness identified the defendant, Roger Harvey, as the robber. Today, I will demonstrate to you that the defendant committed this despicable act, with the testimony of the witness.

**Mr. Reeder (defense attorney):** My client has been caught in an unfortunate and tragic set of circumstances. He was misidentified by the witness, and has been falsely accused of a crime that he did not commit. I will show you the flaws in the prosecution's case, particularly the shortcomings of eyewitness' testimony. After considering everything I have to say today, you should find my client not guilty of the crime with which he has been accused. Thank you.

**DIRECT EXAMINATION OF DR WILLIAMS BY MR. REEDER**

**Mr. Reeder (defence attorney):** Good afternoon, Ma'am. Dr. Williams—could you state and spell your name for the record?

**Dr. Williams:** Carolyn Williams, W-i-l-l-a-m-s.

**Mr. Reeder:** Dr. Williams, what is your occupation?

**Dr. Williams:** I'm a lecturer in psychology.

**Mr. Reeder:** Can you please tell the jury about your educational background?

**Dr. Williams:** I completed my Bachelor of Arts with Honours in Psychology at the University of New South Wales in 2003. In 2004, I began my graduate studies at the University of Melbourne. In 2008 I graduated with my Ph.D. in Psychology.

**Mr. Reeder:** Do you have any areas of specialization in the general area of Psychology?

**Dr. Williams:** My area of specialization is eyewitness identification procedures and eyewitness memory.

**Mr. Reeder:** When we speak of eyewitness identification, tell the jury what you mean?

**Dr. Williams:** In general, we’re talking about individuals’ memories for faces and their ability to later identify a face. We are also talking about people’s reports of what they claim to remember from an event.
Mr. Reeder: This case is about an Armed Robbery that took place on October 23, 2015?
Dr. Williams: Yes, it is.

Mr. Reeder: Were you present at the time?
Dr. Williams: No, I was not.

Mr. Reeder: Have you spoken with the eyewitness?
Dr. Williams: No, I have not.

Mr. Reeder: Are you being paid for your testimony?
Dr. Williams: Yes, I am.

Mr. Reeder: What is your hourly rate?
Dr. Williams: $150/hour.

Mr. Reeder: Does the fact that my office pays you for your opinion in any way influence your expert opinion or findings about human memory and eyewitness identification.
Dr. Williams: No, it does not. My opinion is based on the research, not the specifics of any one case.

Mr. Reeder: Doctor, are you going to give any opinions about the accuracy of the witness in this case?
Dr. Williams: No, I am not.

Mr. Reeder: Is that your role here today?
Dr. Williams: No, that’s not my role.

Mr. Reeder: Why not?
Dr. Williams: My role is to provide the court with information about the scientific research on eyewitness identification and memory. It is up to the jury to decide how to use this information in their consideration of the witness.

Mr. Reeder: Dr. Williams, to start off, can you give the jury a little general information as to how memory works in the context of an eyewitness identification?
Dr. Williams: Our memory is made up of three stages: encoding, retention, and retrieval.

Mr. Reeder: What is encoding?
Dr. Williams: The first stage, encoding, occurs when you are first exposed to the material.

Mr. Reeder: What is retention?
Dr. Williams: The second stage, retention, is the period between when you were first exposed to the material and when you later attempted to retrieve or use that information.

Mr. Reeder: All right, and what is the final stage?
Dr. Williams: The final stage is the retrieval stage, and this is the part where we try to remember or recall what we saw.

Mr. Reeder: Doctor, does your research indicate that there are common misconceptions about how memory works?
Dr. Williams: Yes.

Mr. Reeder: Okay. Can you explain what those are?
Dr. Williams: Some of the misconceptions are that memory is like a video-recorder, that you can essentially play back an event without error. Another misconception is that when a witness or someone views something that was traumatic and memorable, that they are more likely to remember that event. It is true that you are more likely to remember that the traumatic event occurred, but it doesn’t mean that you will remember the specific details of that event.

Mr. Reeder: Doctor, in the field of psychology has there been scientific research conducted on eyewitness accuracy and the factors that affect eyewitness accuracy?
Dr. Williams: Yes.
Mr. Reeder: Before we get into it, can you describe how you research that?
Dr. Williams: The most common form of research conducted in the area of eyewitness identification is experimental research. The primary reason for conducting experimental research is that it gives us the ability to make cause and effect statements, such as “X caused Y.” For example, we could present two equivalent groups of students with a video of a staged crime, and then later test their memory for the event. But, half the students only saw the perpetrator’s face for 2 seconds and the other half saw him for 2 minutes. Well, if we find a difference in their ability to later identify the perpetrator, then we know that the exposure duration (or the length of time that they saw something) affected their memory. In experiments we can also control the ground truth: whether the suspect is guilty or innocent, whereas in real cases we don’t know whether the person who was convicted is in fact the perpetrator.

Mr. Reeder: Why are university students used?
Dr. Williams: University students are commonly used as participants because they are readily available to researchers (who are usually university lecturers). The vast majority of psychological experimental researchers utilize university students in their studies.

Mr. Reeder: Is there any problem with using university students? Could it affect the research?
Dr. Williams: For the most part, no. The process of memory is similar regardless of whether you are a university student or a construction worker. What we learn from experiments is that even under ideal conditions, identifications for people seen briefly, in non-stressful situation, and attempted after brief delays, are frequently inaccurate.

Mr. Reeder: Are there any other types of research?
Dr. Williams: Another way of studying eyewitness accuracy is through examining existing records or data. For example, several archival studies have found that approximately 20% of witnesses in real cases select a filler—a known innocent person—and thus have made an identification error. These results are consistent with results from lab studies, which have found very similar rates of filler selections.

Mr. Reeder: These are real cases right?
Dr. Williams: Yes.

Mr. Reeder: That would sometimes include wrongful convictions?
Dr. Williams: Yes, if someone is convicted based on a mistaken identification, then that would be a wrongful conviction. In fact, in the US alone, DNA evidence has been used to prove that over 337 innocent people have been wrongfully convicted of a crime they did not commit, and a mistaken identification was present in approximately 75% of these cases.

Mr. Reeder: Dr. Williams, you told the jury about the three separate stages of memory?
Dr. Williams: Yes.

Mr. Reeder: As you have researched it, let's go ahead and talk about them. The first one you talked about was the encoding stage?
Dr. Williams: Yes.

Mr. Reeder: Please remind the jury what the encoding phase is?
Dr. Williams: The encoding stage is that first step, when you first see someone.

Mr. Reeder: What factors do scientists say affect memory during the encoding phase?
Dr. Williams: Several factors can influence the information that a witness encodes into memory during the encoding phase, including: how long the witness saw the perpetrator, known as the exposure time; the quality of witnessing conditions (such as lighting, distance); the presence of a weapon, and amount of stress.

Mr. Reeder: Ok, let’s discuss exposure time first. What is exposure time?
**Dr. Williams:** Exposure time refers to how long a person is exposed to the information. In the context of eyewitness identifications, exposure time refers to how long the witness saw the perpetrator’s face.

**Mr. Reeder:** How does it affect accuracy?

**Dr. Williams:** In general, the longer you are exposed to a face, the more likely you are to accurately identify that face.

**Mr. Reeder:** Does it matter if the perpetrator is holding a weapon?

**Dr. Williams:** Yes. In order to encode information, a person has to be paying attention to that piece of information; anything that interferes with your ability to attend to the face will ultimately interfere with your ability to memorize that face. Research has repeatedly shown that a weapon (or any unusual object) captures a person’s attention. As a result, the witness is more likely to pay attention to the weapon itself rather than the face of the individual holding the weapon.

**Mr. Reeder:** Understood. Doctor, what effect does stress or fear have on the ability of an eyewitness to accurately encode information? I think most people would think that when we are stressed we have stronger memories. What do the scientific studies say about that?

**Dr. Williams:** The effect of stress is counter-intuitive. The more stress a person experiences, the less likely they are to remember details of the event. You are less likely to forget that the event occurred, but may not remember specifically what took place, the order of events, or strangers’ faces.

**Mr. Reeder:** Do you know of any studies in this area?

**Dr. Williams:** Stress is difficult to study in the laboratory because our ethical guidelines prohibit us from placing our participants under extremely stressful conditions. In connection with military personnel, however, a group of researchers were able to investigate stress in a mock prisoner of war scenario.

**Mr. Reeder:** What was that?

**Dr. Williams:** In the prisoner-of-war scenario, half of the soldiers were exposed to high stress interrogations and half were exposed to low stress interrogations. Then, one day later, participants were asked to identify their interrogators from a photoboard.

**Mr. Reeder:** How long were they with the interrogators?

**Dr. Williams:** The prisoners were with their interrogators for 40 minutes.

**Mr. Reeder:** What were the results?

**Dr. Williams:** There was a significant decrease in identification accuracy under the high stress. When they were under high stress, 70% of prisoners made a mistake, but under low stress only 50% made a mistake.

**Mr. Reeder:** So, under high stress, only 30% of the prisoners could identify the person who interrogated them for 40 minutes?

**Dr. Williams:** Yes, that’s right.

**Mr. Reeder:** Okay. Now, you talked about the first stage, encoding. We talked about the facts of that. The other two stages are retention and recall?

**Dr. Williams:** Yes.

**Mr. Reeder:** When does the retention phase occur?

**Dr. Williams:** This is the period of time between when you are exposed to the information and when you attempt to recall that information.

**Mr. Reeder:** Does the length of the interval between the event and the identification affect accuracy?
Dr. Williams: Yes, it does. The longer the time between the event and the identification accuracy, the less likely a witness is to correctly select someone from a photoboard. A recent meta-analysis of all the studies in the area demonstrated that our ability to identify someone’s face begins to deteriorate after only 2 hours.

Mr. Reeder: So, it makes sense that our memories get weaker over time. Can our memories actually change over time?

Dr. Williams: Yes.

Mr. Reeder: How do they change?

Dr. Williams: They can change by not only forgetting certain pieces of information, but also by incorporating new information, known as post-event information.

Mr. Reeder: What are some sources of post-event information?

Dr. Williams: Post-event information can come from any number of sources: discussions with other witnesses, police officers, lawyers, as well as any information that one might read in a newspaper or see on tv. The problem is, not everything you hear or read is accurate, but it can still be incorporated into your own memory. Moreover, what we know from that research is that we are relatively bad at remembering the context and the source of the information. The witnesses do not understand that they have acquired post event information, they believe that they saw it at the original event.

Mr. Reeder: Ok, we have talked about the factors that affect when we acquire a memory, and the factors at the retention stage, let's talk about retrieval. When does the retrieval stage happen?

Dr. Williams: The retrieval stage happens when a person tries to remember or recall a piece of information.

Mr. Reeder: In psychological terms, how does somebody retrieve a memory?

Dr. Williams: As I said earlier, we can’t just “replay” the memory. Every time you try to remember something, you are actually rebuilding or reconstructing that memory. It is during this reconstruction that errors can occur because we incorporate information from other sources into our own memory.

Mr. Reeder: Are there any factors that can affect the accuracy of the retrieval process?

Dr. Williams: Yes there are.

Mr. Reeder: What are they?

Dr. Williams: Some of the many factors are how the identification procedure was conducted, what instructions did the officer give prior to conducting a photoboard, and confirmatory post-identification feedback. These factors can affect accuracy at future retrievals, as well as your confidence in that memory.

Mr. Reeder: Let’s discuss an identification procedure. First of all, what is the photoboard?

Dr. Williams: A photoboard consists of the photographs of a suspect and foils and is presented to an eyewitness for identification purposes.

Mr. Reeder: Are there particular studies about photoboards?

Dr. Williams: Yes, photoboards are frequently studied in our area because they are the most common type of identification procedure.

Mr. Reeder: What do those studies suggest about how photoboards relate to accuracy in eyewitness identification?

Dr. Williams: The studies suggest that the way photoboard is presented can have a significant impact on a witness’ identification decision.

Mr. Reeder: How important are instructions given to a witness before the photoboard is presented?
**Dr. Williams:** Instructions given to an eyewitness prior to an identification procedure are of the utmost importance. Suggestive instructions strongly convey to the witness that the perpetrator is in the photoboard and that their job is to pick him out of the photoboard. Ideally, the instructions should state that the perpetrator may or may not be present in the lineup. In the meta-analysis of 15 studies, the non-suggestive instructions reduced identification errors from 70% to 43%.

**Mr. Reeder:** Ok, moving on. What is the difference between a simultaneous and a sequential photoboard?

**Dr. Williams:** A simultaneous photoboard is a traditional photoboard that you have probably seen on tv—all of the photoboard members are presented all at once to the witness. The sequential photoboard was developed in 1985 by Gary Wells and Rod Lindsay, and it has several components to it: (a) the photoboard members are presented one at a time to the witness, (b) the witness must make a yes or no decision for each member before they can move on to the next member, (c) the witness can only see each member once, (d) the witness does not know how many photoboard members they will view, and (e) the person administering the procedure must not know who the suspect is in the photoboard.

**Mr. Reeder:** Is one better than another?

**Dr. Williams:** Yes, the sequential photoboard has a higher accuracy rate than the simultaneous photoboard.

**Mr. Reeder:** Why are sequential photobooks better?

**Dr. Williams:** Witnesses who are exposed to a simultaneous photoboard are thought to engage in a process which is called a relative judgment, in which they compare the photoboard members to each other to decide who looks the most like their memory of the perpetrator. If that person is a close enough match to their memory, then they select that photoboard member as the perpetrator of the crime. This judgment strategy is effective if the suspect is in fact the perpetrator; however, it produces high false identification rates if the perpetrator is not in the photoboard, but is instead replaced with an innocent suspect. The sequential photoboard was designed to reduce reliance on the relative judgment strategy, and instead was designed to encourage witnesses to compare each photoboard member to their memory for the perpetrator and then select that person only if he/she is a good enough match to their memory. Thus, the sequential photoboard has been repeatedly demonstrated to reduce false identification rates compared to the simultaneous photoboard, with only a small reduction in the correct identification rates.

**Mr. Reeder:** During a photoboard procedure, does it matter who does it?

**Dr. Williams:** Yes, it does. An independent person should administer the photoboard procedure—one outside of the investigation who does not know which member is the suspect. This is known as a double-blind procedure.

**Mr. Reeder:** Why?

**Dr. Williams:** This is done to ensure that the administrator doesn’t give any intentional or unintentional cues to the witness as to who they should pick. The witness should also be informed that the administrator has no knowledge of the suspect’s identity so that they do not look to the administrator for cues. So, again, you want the witness to make their decision based on their memory, not based on some cue from the administrator.

**Mr. Reeder:** How does this affect the eyewitness later in court?

**Dr. Williams:** In one study those who made an identification from a non-blind photoboard were significantly more confident in their decision than those who made it from a double-blind photoboard.
Mr. Reeder: Double blind is the best practice?
Dr. Williams: Double-blind is the best practice, not only in eyewitness identification procedures, but also in medical research and psychological research.
Mr. Reeder: Moving on, you’ve also mentioned post-identification feedback, what is it?
Dr. Williams: Post-identification feedback is telling someone that they made the right decision, or that they picked the suspect out of a photoboard.
Mr. Reeder: How does post-identification feedback affect the accuracy of an identification?
Dr. Williams: Post-identification feedback is given after a witness has made an identification, so it doesn’t have an impact on the accuracy of that identification, but it does affect how confident they are in that identification and it also changes their memory for the event itself. Confirmatory comments, for example, “Good, you got the guy!” influence the eyewitness’s confidence in their identification decision, reported quality of witnessing conditions, reported ease with which eyewitness made his or her identification decision, and the amount of attention reportedly paid to the event. This effect is the largest when eyewitness has incorrectly selected an innocent person from the photoboard.
Mr. Reeder: With respect to the effect of post-identification feedback on the eyewitness, again, are people able to put it in context? “I’m more confident because of the feedback I received, versus I’m confident because chose the right person?”
Dr. Williams: No, we generally don’t attribute external influences to our known behavior. We like to think that we're in control of our own opinions and own beliefs, and therefore we're unlikely to say other people influenced us in our memory and our decision.
Mr. Reeder: Have there been any studies about the confidence of witnesses?
Dr. Williams: Nearly every study investigating eyewitness memory has included a measure of the witness’ confidence.
Mr. Reeder: Is there any relationship between how confident a witness is and how accurate they are?
Dr. Williams: We like to think that because a witness is 100% confident in their identification that they must be right, but the research shows us that the relationship between confidence and accuracy is not perfect. Confidence statement given right after the identification decision, prior to any feedback is useful, whereas, confidence during the trial isn’t.
Mr. Reeder: Are there any factors that can affect the relationship between confidence and accuracy?
Dr. Williams: Yes, we talked about post-identification feedback earlier. Telling someone that they chose the right person from the photoboard can further reduce the relationship because now they are extremely confident in their decision, but they aren’t any more accurate.
Mr. Reeder: To sum it up, what are the factors that you believe that a psychologist should tell a lay person to consider in evaluating eyewitness evidence?
Dr. Williams: There are a number of factors that can affect the accuracy of eyewitness’s memory at each of three stages. During the encoding phase these are how long the witness saw the perpetrator, lighting conditions and distance between an eyewitness and the perpetrator, the amount of stress that someone is under, the presence of a weapon.
Mr. Reeder: And what factors during the retention phase should be considered in evaluating accuracy of eyewitness identification?
Dr. Williams: The length of retention time, the opportunity for the witness to have spoken with other witnesses or other people. Any potential leading questions that were given by law
enforcement that might influence the witness's recollection of the events, all those should be considered.

**Mr. Reeder:** Okay, and finally during the retrieval phase what are the factors that should be evaluated?

**Dr. Williams:** You should evaluate how the lineup was presented, what the officer said to eyewitness prior to the identification procedure, type of the photoboard, and whether the officer gave any feedback after eyewitness made an identification.

**Mr. Reeder:** Dr. Williams, are you against the use of photoboards in police procedure?

**Dr. Williams:** No, I'm not against the use of properly conducted photoboards in police procedures.

**Mr. Reeder:** Why not?

**Dr. Williams:** Because a properly conducted photoboard can provide evidence regarding whether or not the police’s suspect is in fact the perpetrator.

**Mr. Reeder:** Dr. Williams, are you against eyewitness identifications being allowed in the justice system?

**Dr. Williams:** No, I am not.

**Mr. Reeder:** Why not?

**Dr. Williams:** Because eyewitness evidence is a common piece of evidence that can be part of the puzzle, but it should not be the only piece of the puzzle that is used to convict someone.

**Mr. Reeder:** Thank you. Nothing further.

---

**CROSS-EXAMINATION OF DR WILLIAMS BY MR. WALKER**

**Mr. Walker (prosecutor):** Good afternoon, Dr. Williams.

**Dr. Williams:** Good afternoon.

**Mr. Walker:** My name is Mr. Walker, by the way. Can we have this understanding: I promise you that I won't be repetitious and waste your time if you can promise me that you won't try to help me by answering questions that I haven't asked, and you'll just answer the questions I did ask. Can we do that?

**Dr. Williams:** Yes.

**Mr. Walker:** Is it true that for many years many judges and many courts didn't let eyewitness experts testify before a jury?

**Dr. Williams:** It is true.

**Mr. Walker:** Okay. And even today, in some jurisdictions in our country and in some courts, eyewitness experts are not allowed to testify in the court?

**Dr. Williams:** It is true.

**Mr. Walker:** Yes ma'am. That's because a lot of people say, a lot of your people, psychologists, psychiatrists say that eyewitness expert testifies about things that are just common sense? Yes or no?

**Dr. Williams:** I think it's misleading to say a "lot of people."

**Mr. Walker:** Reputable people in your field say that, though, don't they?

**Dr. Williams:** A few quite reputable people have expressed that opinion. A few.

**Mr. Walker:** Basically what you testify about is, it's not really hard science, it's more soft science?

**Dr. Williams:** I don't think people in my field would call it that, no, even the very reputable people.
Mr. Walker: During your direct examination with defence counsel, please correct me if I'm wrong, but I don't think I ever heard you mention the defendant’s name, Roger A. Harvey. Did that name ever come out of your mouth?

Dr. Williams: I don't think so.

Mr. Walker: And during your direct examination by defence counsel, correct me if I'm wrong, but I don't think I ever heard you mention the name of the victim?

Dr. Williams: I don't think I mentioned the name, no.

Mr. Walker: Do you know who he is?

Dr. Williams: Yes.

Dr. Williams: Those names didn't just slip your mind, did they?

Dr. Williams: No.

Mr. Walker: Because, of course, you've spoken with all of them because that's why you're here today, to talk about the identification. Have you spoken with them?

Dr. Williams: No. But as it's been indicated, my role today is not to comment on the eyewitness’ accuracy.

Mr. Walker: Like I said, you don't need to help me. My question was: "Have you spoken with them?"

Dr. Williams: No, I haven't.

Mr. Walker: Did you ever ask Mr. Reeder if you could look at his notes from the witness's statements?

Dr. Williams: Well, I have read those. I reviewed all of those. I didn't have to ask him. He sent them to me.

Mr. Walker: He sent those to you?

Dr. Williams: Yes.

Mr. Walker: Okay. Regarding all of those studies that you mentioned, they are just simulations, right? Because you can never exactly recreate a real-life traumatic event like a stabbing or a murder, can you?

Dr. Williams: Well, not ethically, no. That's why we have to look at actual cases and also do the experimental work and look at all of the data.

Mr. Walker: And the experimental work basically boils down to you having some of your university students come in and watch a film and then, after the film, you ask them a bunch of questions.

Dr. Williams: That happens in some of the studies.

Mr. Walker: So, there's no way you can know for sure what witnesses would do in a real-life situation because you can't scientifically recreate that type of event?

Dr. Williams: Well, we certainly can't subject witnesses in the experiments to, you know, a whole lot of trauma that might harm them. That's why we have to do what we can in the experimental studies and complement them with analyses of real-life cases.

Mr. Walker: And the people who criticize your experiments are other psychologists?

Dr. Williams: Sometimes, yes.

Mr. Walker: People whose opinions you respect?

Dr. Williams: Sometimes I do, yes.

Mr. Walker: Okay. Let me ask you a question. Is one factor that increases reliability or accuracy of an eyewitness a short retention interval?

Dr. Williams: Yes.
Mr. Walker: So, if eyewitness makes identification two days after the event, as opposed to, say, two months, the probability that this eyewitness will make accurate identification is higher?

Dr. Williams: Yes.

Mr. Walker: And another factor is the distance between an eyewitness and the perpetrator?

Dr. Williams: Yes.

Mr. Walker: And the less the distance, the more likely you are to be accurate in identifying the person?

Dr. Williams: Generally speaking, yes, but this is not the only factor that is important.

Mr. Walker: Thank you. Nothing further.

Mr. Reeder: Nothing further, your Honor.

CROWN PROSECUTOR CLOSING REMARKS

Mr. Walker (prosecutor): Ladies and gentlemen, after hearing the testimony, I think you'll agree with me that all of the evidence adds up. Eyewitness testimony places Mr Harvey at the scene of the crime. Furthermore, Mr Harvey drives a white ute, which was found not far away from the crime scene. In this ute the police found 8 packs of cigarettes of various brands, black rubbish bag and a blue jumper. The ute was registered to the defendant, who matched witness’s description of the robber. Based on this evidence, I believe it is your duty to find Mr Harvey guilty of an armed robbery.

DEFENSE CLOSING REMARKS

Mr. Reeder (defense attorney): Ladies and gentlemen, I disagree with my colleague’s assessment. The only facts against my client are pieces of circumstantial evidence. The eyewitness was under stress and can’t remember the number plate or provide a detailed description of the robber. My client matches the very vague description that some of you would also match. My client owns a white ute because he is a carpenter. My client smokes a lot, which is why he had several packs of cigarettes. My client did not commit this robbery. For this reason, I feel that you are obligated to find my client not guilty.

JUDGE’S CLOSING REMARKS

Judge Adams: Ladies and gentlemen of the jury, your job is to decide the facts based on the evidence that you heard and saw during the trial. In deciding the facts, you must not be swayed by sympathy or prejudice towards any of the parties. Remember that statements, objections, or arguments made by the lawyers are not evidence in the case. The lawyers' job is to point out those things that are most significant or most helpful to their side of the case. In light of the evidence and the law, do you believe that it is more likely true than not that this robbery was committed by Mr. Harvey? If so, you should decide in favour of the prosecution. On the other hand, if the evidence does not meet the standard or tends to be equally balanced, the prosecution cannot succeed on his or her claim.
Appendix 4. PowerPoint Slides that Accompanied Expert Testimony in Paper 2

Three Memory Stages

- Encoding
- Retention
- Retrieval

Contributing Causes of Wrongful Convictions (first 325 DNA exonerations)
Total is more than 100% because wrongful convictions can have more than one cause.

- Eyewitness Misidentification (235 cases)
- Unvalidated / Improper Forensics (154 cases)
- False Confessions / Admissions (88 cases)
- Informants / Snitches (48 cases)

(Innocence Project, 2016)
Encoding phase

• exposure time
• witnessing conditions
• presence of a weapon
• stress

Prisoner-of-war study

Mistaken Identifications (%)

- High stress
- Low stress

70
50
Retention phase

• Length of interval

• Post-event information
  • Discussions with other witnesses, police officers, lawyers, etc.
  • Media reports

Retrieval stage

• Identification procedure

• Identification instructions

• Post-identification feedback
The bar chart shows the percentage of mistaken identifications under Suggestive Instructions (70%) and Non-suggestive Instructions (43%).

For the Simultaneous vs. Sequential Photoboard, there are two sets of images. The left set, labeled 1-3, is a simultaneous presentation, while the right set, labeled 4-6, is a sequential presentation.

(Source: Carlson & Gronlund, 2011)
Summary

1. Encoding phase
   • how long the witness saw the perpetrator
   • lighting conditions
   • distance
   • amount of stress
   • presence of weapon

Witnesses responding at high end of the scale (Steblay, Wells, & Douglass, 2014)
Summary

2. Retention phase
   • length of retention time
   • discussions with other witnesses
   • leading questions asked by law enforcement personnel

Summary

3. Retrieval phase
   • instructions prior to administration of the photoboard
   • type of photoboard
   • feedback after the identification
Appendix 5. The Eyewitness Instructions and Checklist Transcript

In this case the Crown Prosecutor will present the testimony of the witness who identified the defendant as the person who committed the crime. You must determine whether the witness’s identification of the defendant is reliable and credible, or whether it is based on a mistake or is, for any reason, not worthy of belief.

In this presentation we will discuss the key terms used in Court, I will briefly explain general principles of memory, and describe how the eyewitness researchers conduct the studies. After that I will talk about factors that may affect the eyewitness during the crime, between the crime and the identification procedure, will explain how identification procedure should be conducted, and what factors may influence the eyewitness at the time of the identification.

Terms Used in Court

There are several important terms that will be used throughout the trial. It is essential to be familiar with these words and their definitions.

The culprit is the person who actually committed the crime. If the crime was committed by one person, there is only one culprit. The police, however, can have several suspects during the course of an investigation. When the police have grounds to believe that one of the suspects is the culprit, they can accuse him or her of the crime. During the trial the person accused of committing the crime is called the defendant.

I will now give an overview of the scientific research on eyewitness memory and the factors that are important in determining the reliability of eyewitness evidence. Please pay attention, this information may help you evaluate the reliability of the eyewitness identification evidence.

People can recognize other people from the past experiences and identify them at a later time, but research has shown that people can incorrectly identify an innocent suspect, particularly when the culprit was a stranger. For that reason eyewitness identification of strangers must be given appropriate weight, but it also should be scrutinized carefully to determine the factors that may have affected the likelihood that an eyewitness correctly identified the person who committed the crime—the culprit.

Our memory is not foolproof, it does not act like a video recording device. We cannot replay the memory of a past event exactly as it occurred. Memory is far more complex. The process of remembering consists of three stages: encoding, retention, and retrieval.

Encoding—is the perception of the criminal event. Retention—is the period of time that passes between the event and when they try to recall a piece of information. And retrieval—is the stage during which an eyewitness recalls a criminal event and attempts to identify the culprit. At each of these stages, memory can be affected by a variety of factors.

Informed by the research, I will instruct you on the general factors you should consider when determining whether the eyewitness identification evidence in this case is reliable. Some of these factors may be applicable to the case, whereas others will not be relevant; it is up to you to decide which impairing factors were present and whether they affected the witness.

The types of factors that may affect eyewitness evidence are divided into three groups:

1. Factors that were present at the time of the crime
2. Factors that were present between the criminal event and the identification procedure
3. Factors that were present at the time of the identification

Please note that I will present the factors in chronological order. For example, I describe factors that were present at the time of the crime first because they preceded all other factors, not because they are the most important. The presentation order does not reflect the importance of these factors.

**Please note: Factors that were present at the time of the identification are as important as those present at the time of the crime.**

Before discussing each of these factors, I will briefly explain how psychologists conduct eyewitness studies. The most common form of research conducted in the area of eyewitness identification is experimental laboratory-based research. The primary reason for conducting experimental research is that it gives us the ability to make cause and effect statements, such as “X caused Y.” For example, we could present two equivalent groups of students with a video of a staged crime, and then later test their memory for the event. But, half the students only saw the culprit’s face for 2 seconds and the other half saw him for 2 minutes. Everything else was the same for the two groups except the duration of the crime. If we find a difference in their ability to later identify the culprit, then we know that the length of time that they spent looking at the crime affected their memory. In experiments we can also control the ground truth: whether the suspect is guilty or innocent. This is very important because this is the only type of research in which we can be sure that an eyewitness is accurate or mistaken. The examination of real cases does not provide the same information because we rarely know whether the person who was convicted was in fact the culprit or whether that conviction was a mistake.

The vast majority of eyewitness researchers utilize university students in their studies. Some critics point out that these studies are simulations and that the experience of students who witness a video-recorded crime in the laboratory is not the same as the experience of a real eyewitness or victim. That is true; however, the process of memory is similar regardless of whether we test a university student, an office manager, or a construction worker. What we learn from the experiments is that people can mistakenly identify an innocent suspect even when the identification takes place under non-stressful circumstances and after much shorter delays than those present in real cases.

The following recommendations are based on 50 years of eyewitness research.

In evaluating the reliability of the witness’s identification, you should consider the circumstances present at the time of the crime.

1. **Duration**: The amount of time an eyewitness has to observe an event may affect their ability to identify a stranger. Although there is no minimum time required to make an accurate identification, a brief encounter is less likely to produce an accurate identification compared to prolonged exposure to the culprit. It’s important to understand that a witness’s estimate of how much time they spent with the culprit may not always be accurate because witnesses tend to think events lasted longer than they actually did.

2. **Distance**: A witness may have a better memory for a stranger when they are seen from a short distance than from a large distance. The greater the distance between an eyewitness and the culprit, the higher the risk of a mistake. It’s important to understand that a witness’s estimate of how far he or she was from the culprit may not always be accurate because people tend to have difficulty estimating distances.
(3) **Lighting**: Inadequate lighting may reduce the reliability of an identification because a witness may lack the opportunity to properly see the culprit’s face.

(4) **Presence of a Weapon**: You should consider whether the witness saw a weapon during the crime. A weapon captures a person’s attention. As a result, the witness may be more likely to pay attention to the weapon itself rather than the face of the individual holding the weapon.

(5) **Intoxication**: An identification made by a witness under the influence of a high level of alcohol at the time of the crime tends to be less reliable than an identification by a witness who drank a small amount or no alcohol.

(6) **Disguises**: The culprit’s use of a disguise can affect a witness’s ability to identify the culprit. Disguises like hats, sunglasses, or masks can reduce the accuracy of an identification.

(7) **Stress**: Even under the best viewing conditions, high levels of stress may reduce an eyewitness’s ability to make an accurate identification. The effect of stress on eyewitness’s accuracy is counter-intuitive. The more stress a person experiences, the less likely they are to remember details of the event. They won’t forget that the event occurred, but they may not remember specifically what took place, the order of the events, or accurately remember strangers’ faces.

(8) **Cross-Race Effects**: People have greater difficulty in accurately identifying people from another race because we process other race faces differently than own race faces. When an eyewitness and a culprit are of different races, the probability of a mistaken identification is higher.

In evaluating the reliability of the witness’s identification, you should consider what happened between the criminal event and the identification procedure.

(1) **Delay**: Memories fade with time. As a result, long delays between when the witness saw the crime and when they attempted an identification, may reduce the reliability of the identification.

(2) **Changed Appearance**: If the culprit changes his or her facial features, for example, grows a beard or undergoes surgery, or changes hairstyle between the time of the crime and an identification procedure, this makes the identification more difficult.

Before discussing the third set of factors, it is necessary to talk about some basic principles of identification procedures.

Identification procedures generally take one of two forms. The suspect alone or a single photo of the suspect can be presented to the witness—this procedure is called a showup. Or, the suspect can be presented among several other people who match the description of the suspect—this procedure is called a photoboard.

The most common identification procedure is a photoboard. A photoboard includes photographs of the suspect and “fillers” who are known-innocent people.

Photos in a photoboard can be presented either all at the same time or one after another. Please note that the factors that will be discussed further are relevant regardless of whether the photos were presented one at a time or one after another.

The police creates a photoboard to test their assumption that the suspect is the culprit. Thus, a photoboard acts as a test of this assumption, not as a test of the witness’s memory.

| Just because the suspect is in the photoboard, does not mean that he is guilty. The suspect may not be the culprit. |
The police don’t know whether the suspect is the person who actually committed the crime. But if the witness identifies the suspect, it is taken as the evidence of guilt. It is important to keep in mind, however, that the suspect may be innocent.

If the suspect is the NOT the culprit, any choice made by the witness is mistaken. The only correct decision is to reject the photoboard.

If the suspect IS the culprit, the witness may correctly identify him/her. All other decisions are incorrect.

Further, it is important to draw a distinction between choosing a filler and choosing a suspect. If the witness chooses a filler, he or she has made a mistake. But this mistake has no consequences for the filler, because this person is known to be innocent. On the other hand, if the witness chooses the suspect, but the suspect is not the culprit, this mistake can lead to a prosecution of that innocent suspect. If the witness chooses the suspect who is the culprit, the witness made an accurate identification and it can lead to a prosecution of the culprit.

I will now discuss factors present at the time of the identification.

In evaluating the reliability of the witness’s identification, you should consider the circumstances under which the identification was made.

(1) Showups: A showup is more suggestive than a photoboard, because the suspect is the only possible choice. Nevertheless, it is sometimes necessary for the police to conduct a showup: for example, if the police found a suspect shortly after the crime. Thus, you should take into account the circumstances under which a showup was conducted before deciding whether the risk of using a showup was justified.

(2) Photoboard Composition: The greater the number of choices, the more likely the identification decision will be the result of the witness’s memory and not the result of an improperly conducted identification procedure. A minimum of six persons or photos should be included in the photoboard. The suspect should not stand out from other members of the photoboard.

Here is an example of a fair photoboard in which the suspect does not stand out. The suspect is marked with an arrow for the purposes of this presentation, there should not be any arrows in real photoboards.

Another example illustrates a biased photoboard in which the suspect is clearly different from other members of the photoboard. Please note that this is a sample photoboard, it was not used in the current case.

When a photoboard is biased, witnesses select the suspect more often regardless of his actual guilt. On the other hand, when a photoboard is fair, the suspect is selected more often when he is guilty.

Now I will describe a number of factors that you should consider both if the identification was made from a showup and if the identification was made from a photoboard.

For the sake of simplicity I will describe these factors in relation to a photoboard situation, but please note that the same logic applies to a showup.

(1) Double-blind administration: An administrator who knows the identity of the suspect may intentionally or unintentionally convey that knowledge to the witness, which increases the chance that the witness will identify the suspect, even if the suspect is innocent. For this reason a
photoboard should be administered by a person who does not know the identity of the suspect. Or the officer should not see the photos as the witness looks at them, for example, if it is presented on a computer screen. This double-blind administration ensures that witnesses make their decisions based on their memory, not based on some cue from the administrator, even if that cue was unintentional.

(2) **Instructions**: You should consider what was said to the witness prior to viewing a photoboard. Suggestive instructions strongly convey to the witness that the culprit is in the photoboard and that their job is to pick him out of the photoboard. Instead, identification procedures should begin with non-suggestive instructions to the witness that the culprit *may or may not be* in the photoboard. The witness should not feel compelled to make an identification. Studies have demonstrated that the non-suggestive instructions reduced identification errors from 70% to 43%. Thus, non-suggestive instructions may increase the reliability of eyewitness identification.

(3) **Feedback**: You should consider what was said to the witness after viewing a photoboard. Confirmatory feedback occurs when a police officer or someone else, such as another witness, tells a witness that he or she has correctly identified the culprit. For example: “Good, you got the guy!” However, this statement does not mean that the witness is actually correct, because the suspect may not be the culprit. Feedback is given after a witness has made an identification, so it cannot affect the accuracy of that identification, but it may affect how confident the witness is in that identification and it may also change their memory for the event itself. As this graph illustrates, eyewitnesses who were told that they made the right decision can become more confident that they made an accurate identification, they also reported that they had a better view of the culprit, that it was easier to identify the suspect, and that they paid more attention to the crime.

(4) **Multiple Viewings**: When a witness views the same person more than once, the witness’s memory may become contaminated. For example, a witness may view a photo of the suspect on social media, be later presented with the photo by the detective, and then asked to identify the suspect in a formal photoboard. In this scenario it is impossible to know whether the subsequent identification is the result of the witness’s memory of seeing the culprit committing the crime or if it is the result of the witness’s memory of other instances, like remembering a photo on social media. In other words, a witness may identify the suspect simply because he or she saw this person on a previous occasion, not because they saw the suspect during the crime. If a witness views an innocent suspect in multiple identification procedures, the risk of making a mistake is increased.

(5) **Confidence and Accuracy**: The relationship between confidence and accuracy is not perfect. A witness’s level of confidence during the trial may not be an indication of the accuracy of the identification. The confidence statement given right after the identification decision, but prior to any feedback is useful, whereas confidence during the trial is not. You may learn about the initial confidence statement if it was documented by the police officer or if it was video-recorded.

Please keep in mind that the presence of any impairing factor or combination of factors is not an indication that a particular witness is incorrect. Instead, you should evaluate the factors that I have discussed along with all of the circumstances of this case, including other evidence in this case.
If you determine that the identification was obtained under suggestive circumstances the identification should be afforded no weight. In contrast, if you determine that the identification was obtained under non-suggestive circumstances, you may consider that evidence and decide how much weight to give it. It is up to you to decide the ultimate reliability of the identification.
Appendix 6. The Eyewitness Instructions and Checklist (Checklist)

This checklist is designed to help you evaluate the quality of the eyewitness identification evidence in this case. It is as important to identify factors that deteriorate the quality of identification evidence as it is to identify factors that increase the probability of an accurate identification.

Please remember: the factors are presented in chronological order. The order does not reflect the importance of the factors.

Please complete the following checklist indicating the presence of each factor in the present case.

<table>
<thead>
<tr>
<th>Factors at the time of the crime</th>
<th>Yes</th>
<th>No</th>
<th>Unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Duration.</strong> Did the witness briefly see the culprit?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Distance.</strong> Was the witness far away from the culprit?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Lighting.</strong> Were the lighting conditions poor?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Presence of weapon.</strong> Was a weapon present?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Intoxication.</strong> Was the witness intoxicated?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Disguises.</strong> Did the culprit wear a disguise during the crime?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Stress.</strong> Was the event highly stressful?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cross-race effects.</strong> Are the culprit and the witness of different races?</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Factors between the criminal event and the identification procedure</th>
<th>Yes</th>
<th>No</th>
<th>Unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Delay.</strong> Was there a long delay between the crime and the identification procedure?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Changed appearance.</strong> Did the culprit change appearance between the time of the crime and the identification?</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Factors at the time of the identification</th>
<th>Yes</th>
<th>No</th>
<th>Unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Showup.</strong> Was the defendant identified in a showup (was the defendant's photo shown alone)?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Photoboard composition.</strong> Did the photoboard have fewer than five &quot;fillers&quot; (people other than the defendant)?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did the defendant stand out in the photoboard (other people in the photoboard did not look sufficiently similar to the defendant)?</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Double-blind administration.
Did the photoboard administrator know the identity of the suspect and could see photos while the witness was looking at them?

## Instructions.
The instructions given before the photoboard did not warn the witness that the culprit "May or may not be present in the photoboard".

## Feedback.
Did the photoboard administrator make a confirmatory comment after the witness made the identification (for example: "Good job, that's him/her!", "Yes, that's our suspect!", "Good, you got the guy!", or another variation)?

## Multiple Viewings.
Did the witness view the suspect on multiple occasions (other than during the crime)?

## Confidence.
The witness was not confident during the identification procedure, but appeared highly confident at the trial.

Please remember that it is up to you to decide whether and to what extent these factors affected the witness and to evaluate the identification evidence in combination with all evidence in this case.

**Keeping in mind the presence or absence of the impairing factors listed above, how would you evaluate the overall quality of eyewitness identification evidence in this case? Please explain in your own words.**
Appendix 7. Case Summary and Attorneys’ Arguments used in Paper 3

Case Summary
The following trial concerns the case of Roger A. Harvey who has been charged with the theft of $400 from a wallet of the victim, Sophia Collins. You will hear the detailed description of the crime later.

The incident was reported to the security guard who called the police. The police collected the description of the culprit from the witness. A few hours later the police found a man who matched a general description of the culprit; he had $271 in cash. A week later the witness identified the suspect as the man who committed the crime.

Opening Statement
Judge Adams: The burden of proving the identity of the person who committed the crime is upon the Crown Prosecutor. For you to find this defendant guilty, the Crown Prosecutor must prove beyond a reasonable doubt that this defendant is the person who committed the crime. The defendant has neither the burden nor the duty to show that the crime, if committed, was committed by someone else, or to prove the identity of that other person. You must determine, therefore, not only whether the Crown Prosecutor has proven each and every element of the offence charged beyond a reasonable doubt, but also whether the Crown Prosecutor has proven beyond a reasonable doubt that this defendant is the person who committed it. Reasonable doubt means that if you think there is a reasonable possibility that the accused did not commit the crime, then he is entitled to be found not guilty.

Judge Adams: Ladies and gentlemen of the jury, you will be presented with a case that involves theft. Please keep in mind that the defendant should be presumed innocent until proven guilty. The burden to prove the defendant guilty beyond a reasonable doubt rests solely on the prosecution. The defence does not have to prove anything. Mr Walker, you may begin with your opening statement.

Mr Walker (prosecutor): Ms Collins, the victim, was working at a store when a man stole $400 from her wallet. The defendant matched the description of the culprit and was later identified by the witness as the person who stole the money.

Judge Adams: Mr. Reeder, you may make your opening remarks.

Mr Reeder (defence attorney): My client has been caught in an unfortunate set of circumstances. He was misidentified by the witness, and has been falsely accused of a crime that he did not commit.

Closing Statements
Mr Walker (prosecutor): Ladies and gentlemen, eyewitness testimony places Mr Harvey at the scene of the crime. Moreover, he had a large amount of cash with him when he was arrested not long after the crime. For these reasons, you should find the defendant guilty.

Mr Reeder (defence attorney): I disagree with my colleague’s assessment. My client matches the very vague description that some of you would also match. My client did not commit this robbery and you should find my client not guilty.

Judge Adams: If, after considering all of the evidence, you determine that the Crown Prosecutor has not proven beyond a reasonable doubt that the defendant was the person who committed the crime, then you must find him not guilty. If, on the other hand, after considering all of the evidence, you are convinced beyond a reasonable doubt that the defendant was the person who committed the crime, then you must find him guilty.
Appendix 8. Dependent Variables

1. Do you believe the suspect is guilty? Yes or No [verdict]
2. How confident are you in this decision? 0% = not at all confident, 100% = extremely confident
3. What is the likelihood that that the person identified by the witness is actually the culprit? 0% = very unlikely, 100% = very likely
4. How would you describe the manner in which the lineup was presented to the witness? 0 = completely unfair, 10 = completely fair
5. Perceptions of the witness (combined measure).
   a. Based on what you saw, how good was the witness’ view of the culprit at the time of the crime? 0 = Not good, 10 = Very good
   b. Based on what you saw, how clear was the witness’ memory of the culprit at the time of the lineup? 0 = Not clear, 10 = Very clear
   c. Based on what you saw, how much attention do you think the witness paid to the culprit’s face at the time of the crime? 0 = None, 10 = A lot
6. Perceptions of the lineup administrator (combined measure).
   a. Did the officer pressure the witness to choose someone from the lineup? 0 = No pressure, 10 = Definite pressure
   b. Did the officer influence the witness’ confidence in their decision? 0 = No influence, 10 = Definite influence
   c. Based on what you saw, at the time of the lineup procedure, do you think the officer knew which lineup member was the suspect? 0 = Definitely did not know, 10 = Definitely knew
   d. To what extent do you think the officer was invested in the outcome of the lineup? 0 = Not invested in the outcome, 10 = Extremely invested in the outcome

Questionnaire: Eyewitness Factors

1. How many stages of memory are involved in the process of making a positive identification?
   a. 1
   b. 3
   c. 5
2. What are the stages of memory that are involved in the process of making a positive identification?
   a. Observation; Questioning; Answering
   b. Action; Questioning; Identification
   c. Observation; Retention; Identification
   d. Identification evidence in potentially unreliable.
3. Identification evidence in potentially unreliable.
   a. True
   b. False
4. A witness may think that s/he is accurate and still be mistaken.
5. What are the three factors that were discussed in the instructions that are relevant to assessing the reliability of a witness’ identification?
   a. Lighting conditions; stressfulness of the event; gender of lineup administrator
   b. Circumstances of observation; the characteristics of the witness; factors concerning the identification
   c. An attitude of the eyewitness; cross-race identification; number of corroborating witnesses

6. Which of these factors should not be considered when evaluating the circumstances of observation?
   a. Eyewitness’s mood
   b. Distance between the eyewitness and the perpetrator
   c. Distinct features of the perpetrator

7. Which of these factors should not be considered when evaluating the characteristics of the witness?
   a. The quality of the witness as an observer
   b. Whether the eyewitness was tired
   c. Gender of the eyewitness

8. Stress increases the accuracy of a witness’ identification.
   a. True
   b. False

9. An identification procedure is fair if:
   a. Other people in the photoboard did not look similar to the defendant
   b. The police officer did not know who the suspect is
   c. The eyewitness saw the photo of the defendant before attempting to identify the culprit

10. The degree of similarity between the suspect and other people in the photoboard can influence accuracy of identification.
    a. True
    b. False
Appendix 9. Ethics Declaration

All conditions pertaining to the clearances (attached) were properly met. Annual progress reports and final reports have been submitted to the Swinburne Human Research Ethics Committee (SUHREC).

Signature: Alena Skalon, 17th March 2019
Appendix 10. Ethics Approval Statements

Papers 1 and 3 Human Research Ethics Approval

Paper 1 Study Addition to an Existing Project

Alena Skalon

From: Jennifer Beaudry
Sent: Monday, 11 April 2016 10:06 AM
To: Alena Skalon
Subject: FW: SUHREC Project 2013/052 Ethics Modification/Extension (10)

From: Sally Fried On Behalf Of RES Ethics
Sent: Monday, 11 April 2016 10:03 AM
To: RES Ethics; Jennifer Beaudry
Subject: SUHREC Project 2013/052 Ethics Modification/Extension (10)

To: Dr Jennifer Beaudry, FHAD

Dear Dr Beaudry,

SUHREC Project 2013/052 Does camera perspective bias evaluations of eyewitness identification evidence?
Dr Jennifer Beaudry, FHAD
Approved Duration: 07/05/2013 to 07/05/2016; extended to 31/12/2017.

I refer to your e-mail of 8 April 2016 in which you requested the addition of a study; change in consent instrumentation (Consent Information Statement and Debriefing Form); and project personnel changes (change in institution of co-investigator Dr James Sauer; removal of investigators Nimesha Fernando, Mai Tram Nguyen, and Caitlin Reed; and change in role of Alena Skalon from ‘other project personnel’ to student Investigator). The documentation was reviewed by a SHESCI delegate.

I am pleased to advise that, as modified to date, the project/protocol may continue in line with standard ethics clearance conditions previously communicated and reprinted below.

Please contact me if you have any queries about on-going ethics clearance, citing the SUHREC project number. Copies of clearance emails should be retained as part of project record-keeping.

As before, best wishes for the project.

Kind regards,

Sally Fried

Secretary, SHESCI
Dear Dr Beaudry,

SUHREC Project 2013/052 The Trial (Formerly ‘Does camera perspective bias evaluations of eyewitness identification evidence?’)

Dr Jennifer Beaudry, FHAD
Approved Duration: 07/05/2013 to 07/05/2016; extended to 31/12/2017.

I refer to Alena Skalon’s e-mail of 14 March 2017 in which she requested a modification to the project by the addition of a study; removal of Associate Investigator Mitchell Kloet and Research Assistant Melinda Carboon; addition of student investigators Bonnie Kirkman and Travis Edmonds; addition of Research Assistant Ollivia Bradfield, and change in Project Title to ‘The Trial’. The documentation was reviewed by a SHESC1 delegate.

I am pleased to advise that, as modified to date, the project/protocol may continue in line with standard ethics clearance conditions previously communicated and reprinted below.

Please contact me if you have any queries about on-going ethics clearance, citing the SUHREC project number. Copies of clearance e-mails should be retained as part of project record-keeping.

As before, best wishes for the project.

Kind regards,

Sally Fried

Secretary, SHESC1
Paper 2, Study 1. Ethics Approval

Aiena Skalon

From: Astrid Nordmann
Sent: Wednesday, 18 May 2016 1:52 PM
To: Jennifer Beaudry
Cc: RES Ethics; Aiena Skalon
Subject: SHR Project 2016/087 - Ethics clearance

To: Dr Jennifer Beaudry, FHAD

Dear Jennifer,

SHR Project 2016/087 – Eyewitness Evidence
Dr Jennifer Beaudry, Aiena Skalon [Student], Jamie Smith-Morvell [Student], Mr Mitchell Kloet [Student], Ms Sophie Simmonds [Student], Ms Meg Blackie [Student], Melinda Carboon [Student] – FHAD
Approved duration: 18-05-2016 to 17-01-2020 [adjusted]

I refer to the ethical review of the above project by a Subcommittee (SHESC1) of Swinburne’s Human Research Ethics Committee (SUHREC). Your responses to the review as emailed on 18 May 2016, were put to the Subcommittee delegate for consideration.

I am pleased to advise that, as submitted to date, ethics clearance has been given for the above project to proceed in line with standard on-going ethics clearance conditions outlined below.

- All human research activity undertaken under Swinburne auspices must conform to Swinburne and external regulatory standards, including the National Statement on Ethical Conduct in Human Research and with respect to secure data use, retention and disposal.

- The named Swinburne Chief Investigator/Supervisor remains responsible for any personnel appointed to or associated with the project being made aware of ethics clearance conditions, including research and consent procedures or instruments approved. Any change in chief investigator supervisor requires timely notification and SUHREC endorsement.

- The above project has been approved as submitted for ethical review by or on behalf of SUHREC. Amendments to approved procedures or instruments ordinarily require prior ethical appraisal/clearance. SUHREC must be notified immediately or as soon as possible thereafter of (a) any serious or unexpected adverse effects on participants and any redress measures; (b) proposed changes in protocols; and (c) unforeseen events which might affect continued ethical acceptability of the project.

- At a minimum, an annual report on the progress of the project is required as well as at the conclusion (or abandonment) of the project. Information on project monitoring and variations/additions, self-audits and progress reports can be found on the Research Intranet pages.

- A duly authorised external or internal audit of the project may be undertaken at any time.

Please contact the Research Ethics Office if you have any queries about on-going ethics clearance, citing the Swinburne project number. A copy of this email should be retained as part of project record-keeping.
Paper 2, Study 2. Ethics Approval

Alena Skalon

From: Sally Fried on behalf of RES Ethics
Sent: Wednesday, 8 March 2017 2:53 PM
To: Jennifer Beaudry
Cc: RES Ethics, Alena Skalon
Subject: SHR Project 2016/087 - Ethics Extension/Modification (1)

To: Dr Jennifer Beaudry, FHAD

Dear Jennifer,

SHR Project 2016/087 – Jurors’ Perceptions of Eyewitness Evidence (Formerly ‘Eyewitness Evidence’)
Dr Jennifer Beaudry, Alena Skalon (Student), Jamie Smith-Morvell (Student), Ms Sophie Simmonds (Student), Ms Meg Blackle (Student), Travis Edmonds (Student) – FHAD
Approved duration: 18-05-2016 to 17-01-2020 [adjusted]
Modified: March 2017

I refer to your e-mail of 8 March 2017 in which you requested a modification to the project by the addition of a study; change of project title to ‘Jurors’ Perceptions of Eyewitness Evidence’; addition of student investigator Travis Edmonds; and, removal of student investigators Mitchell Kloet and Melinda Carboon. The documentation was reviewed by a SHESC1 delegate.

I am pleased to advise that, as modified to date, the project/protocol may continue in line with standard ethics clearance conditions previously communicated and reprinted below.

Please contact me if you have any queries about on-going ethics clearance, citing the SUHREC project number. Copies of clearance e-mails should be retained as part of project record-keeping.

As before, best wishes for the project.

Kind regards,

Sally Fried

Secretary, SHESC1
Appendix 11. Ethics Final Reports

Papers 1 and 3. Acknowledgement of the Report Receipt

Alena Skalon

From: Jennifer Beaudry
Sent: Monday, 4 March 2019 1:14 PM
To: Alena Skalon
Subject: FW: Acknowledgement of Report for SUHREC Project - 2013/052

On 1/3/19, 2:14 pm, "resethics@swin.edu.au" <resethics@swin.edu.au> wrote:

Dear Jennifer,

Re: End of Student Involvement Report for the project 2013/052

'The Trial (Formerly 'Evaluations of Eyewitness Identification Evidence')' (Report Date: 01-03-2019)

The End of Student Involvement report for the above project has been processed and satisfies the reporting requirements set under the terms of ethics clearance.

Thank you for your attention to this matter.

Regards
Research Ethics Team

Swinburne Research (H68)
Swinburne University of Technology
PO Box 218
HAWTHORN VIC 3122
Tel: 03 9214 3845
Fax: 03 9214 5267
Email: resethics@swin.edu.au
Paper 2. Acknowledgement of the Report Receipt

Alena Skalon

From:          Jennifer Beaudry
Sent:          Thursday, 14 March 2019 3:57 PM
To:            Alena Skalon
Subject:       FW: Acknowledgement of Report for SUHREC Project - 2016/087

On 14/3/19, 3:45 pm, "resethics@swin.edu.au" <resethics@swin.edu.au> wrote:

Dear Jennifer,

Re: Final Report for the project 2016/087

‘Jurors’ Perceptions of Eyewitness Evidence (Formerly ‘Eyewitness Evidence’) (Report Date: 14-03-2019)

The Final report for the above project has been processed and satisfies the reporting requirements set under the terms of ethics clearance.

Thank you for your attention to this matter.

Regards
Research Ethics Team

Swinburne Research (H68)
Swinburne University of Technology
PO Box 218
HAWTHORN VIC 3122
Tel: 03 9214 3845
Fax: 03 9214 5267
Email: resethics@swin.edu.au
Appendix 12. List of Presentations and Publications

Conference Presentations


Publications

